DESIGN OF EXPERIMENTS FOR MULTIRESPONSE MODELS

By

MANEL COORAY WIJESINHA

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN
PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

1984

ACKNOWLEDGMENTS

I thank the chairman of my supervisory committee, Professor Andre I. Khuri, for his guidance and support throughout the writing of this dissertation.

I thank the other members of my committee, Professors John. A. Cornell, Esam M. Ahmed, Victor Chew, and Kenneth P. Portier for their assistance, and also Professor P. V. Rao who has been very helpful to me.

I thank Carole Boone for her careful and efficient typing of this manuscript.

I take this opportunity to express my special gratitude to my parents, sister, and two brothers for their constant encouragement and support given to me throughout the years I was a student.

Finally, I acknowledge with love and appreciation everything that my husband, Alex, has done for me.

TABLE OF CONTENTS

Abstract of Dissertation Presented to the Graduate School of the University of Florida in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

DESIGN OF EXPERIMENTS FOR MULTIRESPONSE MODELS

By

MANEL COORAY WIJESINHA

December 1984

Chairman: Andre I. Khuri
Major Department: Statistics

We deal with the construction of designs for linear multiresponse models. Three separate topics in this area are considered.

The first topic is the generation of multiresponse D-optimal designs when the variance-covariance matrix of the random error vector is not known. If $E(Y_i) = F_i\underline{\beta}_i$, $i = 1, 2, \ldots, r$ are the fitted models in a multiresponse experiment with r responses, then a D-optimal design consists of a set of design points $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_N$ which maximizes the determinant of $F'(\Sigma^{-1}\otimes I_N)F/N$ where F is the block diagonal matrix of the form $F = \text{diag}(F_1, F_2, \ldots, F_r)$ and Σ is the variance-covariance matrix of the random error vector. The design points are chosen iteratively in a process which involves the estimation of Σ by a consistent estimator at every iteration.

The second topic concerns the development of two design criteria to improve the power of the multiresponse lack of fit test associated with

a linear multiresponse model. These design criteria are multivariate extensions of the Λ_1 and Λ_2-optimality criteria for the single response case discussed by E. R. Jones and T. J. Mitchell (Design criteria for detecting model inadequacy, Biometrika, 65 [1978], pp. 541-551). A procedure is presented for the sequential generation of a design based on the Λ_2-criterion.

The third topic is the construction of robust designs to reduce the effect of nonnormality of the error distribution on tests of hypotheses associated with linear combinations of the parameter vectors in a linear multiresponse model.

Each of the above topics is important in the area of multiresponse surface designs. Furthermore, the techniques we develop could be used for better design of multiresponse experiments.

CHAPTER ONE
INTRODUCTION

1.1 A Note on Multiresponse Surface Methodology

An experiment in two or more responses are measured for each setting of a group of independent variables is called a multiresponse experiment. As an example, an industrial engineer may want to study the influence of cutting speed and depth of cut on the life of a tool and the rate at which it loses metal. Similarly, a medical researcher studying the effects of complexing agents on the yield of a certain antibiotic may also be interested in the production cost, and a food scientist may wish to determine how the flakiness, gumminess, and specific volume of a pie crust vary according to the content of water, flour and shortening. The settings of the independent variables are usually controlled by the experimenter and these variables are therefore called controllable variables.

Perhaps the most important objective of the analysis of such data is the estimation of the relationships between the responses themselves as well as between the responses and the controllable variables within some specified region of interest. If there are r responses y_1, y_2, ..., y_r and k controllable variables ξ_1, ξ_2, ..., ξ_k, these relationships, or models, can be written as

$$y_{ui} = \eta_i(\underline{\xi}_u, \underline{\beta}_i) + \varepsilon_{ui}, \quad i=1, 2, \ldots, r; \quad u=1, 2, \ldots, N \tag{1.1}$$

where y_{ui} is the u^{th} observation on the i^{th} response, $\underline{\xi}_u = (\xi_{u1}, \xi_{u2}, \ldots, \xi_{uk})'$ is the vector of controllable variables at the u^{th} setting, $\underline{\beta}_i$, is the vector of unknown parameters associated with the i^{th} response, and ε_{ui} is the experimental error in the u^{th} observation on the i^{th} response ($i=1, 2, \ldots, r$; $u=1, 2, \ldots, N$). In the development which follows we will find it convenient to work with unitless variables instead of the independent (previously called controllable) variables. We define the unitless variables in coded form as

$$x_{ui} = \frac{\xi_{ui} - \bar{\xi}_i}{S_i} \quad i=1, 2, \ldots, r; \quad u=1, 2, \ldots, N$$

where $\bar{\xi}_i$ is the average of ξ_{ui} and S_i is some scale constant. Hereafter we refer to the x_{ui} as controllable variables.

A vector-valued function consisting of the r responses, $y_1, y_2, \ldots, y_r$, is called a multiresponse function, and the univariate models given in (1.1) make up the corresponding multiresponse model. The exact forms of the η_i's are not usually known, and will be approximated by polynomials which are chosen by the experimenter in advance. Observations on all r responses can then be collected at some specific values of $\underline{x} = (x_1, x_2, \ldots, x_k)'$ determined by a suitable design. These values of $\underline{x}$ are called design points and the choice of the design is based on an appropriate criterion which depends on the model given in

(1.1). After the data have been collected, the unknown parameter vectors $\underline{\beta}_i$, i=1, 2, ..., r are estimated using a technique such as least squares or maximum likelihood estimation and the fitted response models become completely determined. These models can be used to

1. predict the response values at any given setting of the controllable variables within the experimental region
2. determine the optimum conditions on the set of controllable variables that optimize the multiresponse function
3. test hypotheses concerning the multiresponse model's parameters.

The predicted response functions can provide a good representation of the behavior of the true responses depending on the adequacy of the fitted model forms and the precision with which the parameters are estimated. Thus in most situations the design criteria are intended to improve the precision of the parameter estimates and to increase the power of tests associated with detecting model inadequacies.

1.2 Literature Review

Aitken (1935) suggested that the method of generalized least squares could be applied to estimate the parameters associated with a multiresponse model. Aitken's estimators, however, depend on the variance-covariance matrix Σ associated with the error vector which is usually unknown. Zellner (1962) obtained estimates for the elements of Σ based on the residuals derived from an equation-by-equation application of ordinary least squares. This estimation procedure is

briefly explained in Chapter Two and it was shown by Zellner that the parameter estimates thus obtained were asymptotically more efficient than single equation least squares estimators. A more general method of multiresponse parameter estimation which applies to nonlinear as well as linear models was developed by Box and Draper (1965), by following a Bayesian approach using non-informative prior distributions for all the parameters in the multiresponse model, including the elements of Σ. The estimates for $\underline{\beta} = (\underline{\beta}_1, \underline{\beta}_2, \ldots, \underline{\beta}_r)'$, were obtained by minimizing the determinant of the rxr matrix $V = WW'$ with respect to $\underline{\beta}$, where

$$W' = [w_{ui}] = [y_{ui} - \eta_{ui}] \quad ,$$

and (1.2)

$$\eta_{ui} = \eta_i(\underline{x}_u, \underline{\beta}_i), \qquad i = 1, 2, \ldots, r; \quad u = 1, 2, \ldots N \quad .$$

The matrix V is called the dispersion matrix. This estimation procedure does not require knowledge of Σ, thus eliminating the need for its estimation.

It was, however, shown by Box et al. (1973) that this estimation procedure can be adversely affected by the presence of linear dependencies in the data. Therefore it is important that such dependencies be removed prior to the maximization of the determinant of V. In some cases the experimenter may not even be aware of these dependencies. Box et al. (1973) gave a procedure to detect these relationships among the data by means of an eigenvalue-eigenvector analysis based on the matrix DD' where

$$D' = [d_{ui}] = [y_{ui} - \bar{y}_i], \quad i = 1, 2, \ldots, r; \quad u = 1, 2, \ldots, N \ .$$

McLean et al. (1979) indicated that the presence of singularities in V depended not only on the data but also on the form of the model. In other words, inadequacies in the fitted models may also jeopardize the technique of parameter estimation. It is thus quite desirable to perform a test of multiresponse lack of fit to detect any inadequacies in the fitted models. Khuri (1983) developed a multiresponse lack of fit test and also a procedure that can be used to determine those responses which are most contributors to lack of fit. A brief discussion of this test is given in Chapter Three.

In past statistical literature the topic of multiresponse designs has been largely ignored except for the work done by Fedorov (1972, Ch. 5) on multiresponse D-optimal designs when the variance-covariance matrix of the error vector is known.

1.3 Objectives of the Present Work

The main purpose of this research is to construct designs for linear multiresponse models. Three separate topics in this area are considered. In Chapter Two, a sequential procedure is developed to construct D-optimal designs when the variance-covariance matrix of the error vector is not known. In addition, a convergence proof for this procedure is given. In Chapter Three, two design criteria are developed to improve the power of the multiresponse lack of fit test. These are the multivariate extensions of the Λ_1 and Λ_2 optimality criteria

defined by Jones and Mitchell (1978). An iterative design algorithm which converges to a Λ_2-optimal design is also presented. In Chapter Four, the construction of designs to reduce the effect of nonnormality of the error distribution on tests of hypotheses (robust multiresponse designs) is discussed. These tests deal with linear combinations of the parameter vectors and the design criterion is to minimize a multivariate measure of kurtosis for the controllable variables, derived by Mardia (1971). The final chapter consists of concluding remarks and suggestions for future research.

CHAPTER TWO
D-OPTIMAL DESIGNS

2.1 Introduction

There has been a great deal of attention focussed on the topic of D-optimal designs, especially in single response experiments. The main importance of such designs is that they minimize the volume of the confidence region of the regression parameter vector and thus increase the precision of the parameter estimates. St. John and Draper (1975) have written an extensive survey paper on this subject that traces its historic development and contains a comprehensive bibliography. The paper also discusses algorithms to construct D-optimal designs including the well known ones developed by Fedorov (1972) and Wynn (1970).

Construction of D-optimal designs for multiresponse experiments was considered by Fedorov (1972). However, his algorithm required that the variance-covariance matrix, Σ, of the error vector be known. This is rarely the case in practice. The primary objective of this chapter is to develop an algorithm to construct multiresponse D-optimal designs when Σ is not known.

2.2 Model and Notation

In this section we introduce our notation for this chapter. We consider a situation in which observations are made on r responses at N experimental runs (not necessarily all distinct) such that the i^{th} response at the u^{th} experimental run obeys the model

$$y_{ui} = \eta_i(\underline{x}_u, \underline{\beta}_i) + \varepsilon_{ui}, \quad i = 1, 2, \ldots, r; \quad u = 1, 2, \ldots, N \quad . \tag{2.1}$$

Here y_{ui} is the observation on the i^{th} response at the u^{th} setting, $\underline{x}_u = (x_{u1}, x_{u2}, \ldots, x_{uk})'$ is the vector of controllable variables at the u^{th} setting and belongs to the experimental region χ; a compact subset in the k-dimensional Euclidean space, $\underline{\beta}_i$, i = 1, 2, ..., r is a vector of p_i unknown parameters associated with the i^{th} response model, and ε_{ui} is the random error in the i^{th} response for the u^{th} run. The set of points $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_N$ forms an N-point discrete design and is denoted by D_N. Assuming that every model in (2.1) is linear in the parameters, we can write (2.1) as

$$y_{ui} = \underline{f}_i'(\underline{x}_u)\underline{\beta}_i + \varepsilon_{ui}, \quad i = 1, 2, \ldots, r; \quad u = 1, 2, \ldots, N \tag{2.2}$$

where $\underline{f}_i(\underline{x})$ is a p_i x 1 vector which depends on the form of the response function assumed by the i^{th} model and whose elements are known functions

of the controllable variables assumed continuous within the experimental region χ. Using matrix notation the above model can be written as

$$\underline{Y}_i = F_i \underline{\beta}_i + \underline{\varepsilon}_i \qquad i = 1, 2, \ldots, r \tag{2.3}$$

where $\underline{Y}_i$ is an Nx1 vector of observations on the i^{th} response, F_i is an Np_i matrix of full column rank with the u^{th} row containing the elements of $\underline{f}_i'(\underline{x}_u)$, $u = 1, 2, \ldots, N$, and $\underline{\varepsilon}_i$ is an Nx1 vector of random errors associated with the ith response ($i = 1, 2, \ldots, r$), with $E(\underline{\varepsilon}_i) = 0$ and $E(\underline{\varepsilon}_i \underline{\varepsilon}_j') = \sigma_{ij} I_N$; $i, j = 1, 2, \ldots, r$. This implies that the errors associated with any single response are homoscedastic and uncorrelated. Thus the variance-covariance matrix of the Nrx1 vector $\underline{\varepsilon} = (\underline{\varepsilon}_1', \underline{\varepsilon}_2', \ldots, \underline{\varepsilon}_r')'$ is given by

$$\operatorname{Var}(\underline{\varepsilon}) = \begin{bmatrix} \sigma_{11}I_N & \sigma_{12}I_N & \cdots & \sigma_{1r}I_N \\ \sigma_{21}I_N & \sigma_{22}I_N & \cdots & \sigma_{2r}I_N \\ \vdots & & & \vdots \\ \sigma_{r1}I_N & \sigma_{r2}I_N & \cdots & \sigma_{rr}I_N \end{bmatrix} = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1r} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2r} \\ \vdots & & & \vdots \\ \sigma_{r1} & \sigma_{r2} & \cdots & \sigma_{rr} \end{bmatrix} \otimes I_N \tag{2.4}$$

$$= \Sigma \otimes I_N$$

where I_N is the identity matrix of order NxN and '$\otimes$' denotes the direct product. The system of equations in (2.3) may also be written as

$$\begin{bmatrix} \underline{Y}_1 \\ \underline{Y}_2 \\ \vdots \\ \underline{Y}_r \end{bmatrix} = \begin{bmatrix} F_1 & 0 & \cdots & 0 \\ 0 & F_2 & \cdots & 0 \\ & & & \\ 0 & & \cdots & F_r \end{bmatrix} \begin{bmatrix} \underline{\beta}_1 \\ \underline{\beta}_2 \\ \vdots \\ \underline{\beta}_r \end{bmatrix} + \begin{bmatrix} \underline{\varepsilon}_1 \\ \underline{\varepsilon}_2 \\ \vdots \\ \underline{\varepsilon}_r \end{bmatrix} \qquad (2.5)$$

or

$$Y = F_{D_N} \underline{\beta} + \underline{\varepsilon}$$

where $Y' = [\underline{Y}_1', \underline{Y}_2', \ldots, \underline{Y}_r']$, $\underline{\beta}' = [\underline{\beta}_1', \underline{\beta}_2', \ldots, \underline{\beta}_r']$, $\underline{\varepsilon}' = [\underline{\varepsilon}_1', \underline{\varepsilon}_2', \ldots, \underline{\varepsilon}_r']$, and F_{D_N} represents the block diagonal matrix on the right hand side of (2.5). Let p be the total number of unknown parameters in the system. Then $p = \sum_{i=1}^{r} p_i$ is the dimension of $\underline{\beta}$. Since each F_i is of full column rank, F_{D_N} is also of full column rank equal to p. As suggested by Aitken (1935) the technique of generalized least squares can be employed to obtain an estimate for $\underline{\beta}$ given by

$$\hat{\underline{\beta}} = [F_{D_N}'(\Sigma^{-1} \otimes I_N) F_{D_N}]^{-1} F_{D_N}'(\Sigma^{-1} \otimes I_N) \underline{Y} \quad , \qquad (2.6)$$

and the variance-covariance matrix of $\hat{\underline{\beta}}$ is

$$\mathrm{Var}(\hat{\underline{\beta}}) = [F'_{D_N}(\Sigma^{-1} \otimes I_N)F_{D_N}]^{-1} \quad . \tag{2.7}$$

Then at point $\underline{x}$ the predicted response vector is

$$\begin{bmatrix} \hat{y}_1(\underline{x}) \\ \hat{y}_2(\underline{x}) \\ \vdots \\ \hat{y}_r(\underline{x}) \end{bmatrix} = \begin{bmatrix} \underline{f}'_1(\underline{x}) & 0 & \cdots & 0 \\ 0 & \underline{f}'_2(\underline{x}) & \cdots & 0 \\ & & & \\ 0 & & \cdots & \underline{f}'_r(\underline{x}) \end{bmatrix}_{r\mathrm{x}p} \hat{\underline{\beta}} \tag{2.8}$$

or

$$\hat{\underline{y}}(\underline{x}) = \phi'(\underline{x})\hat{\underline{\beta}}$$

with variance-covariance matrix

$$\mathrm{Var}[\hat{\underline{y}}(\underline{x})] = \phi'(\underline{x})[F'_{D_N}(\Sigma^{-1} \otimes I_N)F_{D_N}]^{-1}\phi(\underline{x}) \tag{2.9}$$

where $\hat{\underline{y}}(\underline{x}) = (\hat{y}_1(\underline{x}), \hat{y}_2(\underline{x}), \ldots, \hat{y}_r(\underline{x}))'$ and $\phi'(\underline{x})$ represents the rxp block diagonal matrix on the right hand side of (2.8).

2.3 Design Theory

The concept of a design measure defined below is important to the theory of optimal designs.

Definition 2.1

A design measure defined on an experimental region χ is a probability measure, $\zeta(\underline{x})$ on χ, which satisfies

$$\zeta(\underline{x}) \geq 0 \quad \text{and} \quad \int_{\chi} d\zeta(\underline{x}) = 1 \quad , \quad \underline{x} \in \chi \qquad (2.10)$$

(see Wynn (1970), Kiefer (1959), Fedorov (1972)).

Given a measure $\zeta(\underline{x})$ on χ, the set of points $\underline{x} \in \chi$ at which $\zeta(\underline{x}) > 0$ is called the support or spectrum of $\zeta(\underline{x})$. An example of a design measure is as follows. Let $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_s \in \chi$ and let $\zeta(\underline{x})$ defined as

$$\zeta(\underline{x}) = \begin{cases} 0 & \underline{x} \neq \underline{x}_i \qquad i = 1, 2, \ldots, s \\ \lambda_i & \underline{x} = \underline{x}_i \quad , \quad i = 1, 2, \ldots, s \end{cases}$$

where $\sum_{i=1}^{s} \lambda_i = 1$ with $0 \leq \lambda_i \leq 1$. The design measure $\zeta(\underline{x})$ is said to be discrete if λ_i is a rational number for all $i = 1, 2, \ldots, s$, otherwise, if at least one λ_i is irrational, then it is continuous. In particular, we define the discrete design measure $\zeta_{D_N}(\underline{x})$ by

$$\zeta_{D_N}(\underline{x}) = \begin{cases} 0 & \underline{x} \neq \underline{x}_u \quad u = 1, 2, \ldots, s \\ \frac{n_u}{N} & \underline{x} = \underline{x}_u \ , \quad u = 1, 2, \ldots, s \end{cases}$$

where n_u is the number of replications at the point $\underline{x}_u$ and $\sum_{u=1}^{s} n_u = N$. In this manner we can associate a design measure with any given design. For a given design measure $\zeta(\underline{x})$ on χ and a given variance-covariance matrix Σ, we define the moment matrix $M(\zeta, \Sigma)$ as

$$M(\zeta, \Sigma) = \int_{\chi} \phi(\underline{x})\Sigma^{-1}\phi'(\underline{x})d\zeta(\underline{x}) \quad . \tag{2.11}$$

In particular, for a discrete design measure $\zeta_{D_N}(\underline{x})$

$$M(\zeta_{D_N}, \Sigma) = \frac{1}{N} \sum_{u=1}^{N} \phi(\underline{x}_u)\Sigma^{-1}\phi'(\underline{x}_u) \tag{2.12}$$

or equivalently

$$M(\zeta_{D_N}, \Sigma) = \frac{F'_{D_N}(\Sigma^{-1} \otimes I_N)F_{D_N}}{N} \quad . \tag{2.13}$$

2.3.1 Properties of $M(\zeta, \Sigma)$

For proofs see Fedorov (1972, p. 210).

Let H be the class of all design measures on χ. Then

1. For $\zeta \in H$, $M(\zeta, \Sigma)$ is positive semidefinite.

2. The set $\Lambda(\Sigma)$ of all matrices $M(\zeta, \Sigma)$, $\zeta \in H$ is convex i.e., if $0 \leq \alpha \leq 1$ and $M_1, M_2 \in \Lambda(\Sigma)$, then $M = (1 - \alpha)M_1 + \alpha M_2 \in \Lambda(\Sigma)$.

3. For $\zeta \in H$, $M(\zeta, \Sigma)$ can be represented as

$$M(\zeta, \Sigma) = \sum_{u=1}^{s} \lambda_u \phi(\underline{x}_u)\Sigma^{-1}\phi'(\underline{x}_u) \quad , \tag{2.14}$$

where $1 \leq s \leq p'$, $0 \leq \lambda_u \leq 1$ with $\sum_{u=1}^{s} \lambda_u = 1$, and $p' = (p(p + 1)/2) + 1$.

In view of (3) it is clear that given any design measure ζ, there exists a design measure ζ' which has a finite support with at most p' distinct points such that $M(\zeta, \Sigma) = M(\zeta', \Sigma)$. However the fact that ζ' has finite support does not mean that ζ' is associated with a discrete design, since the value of $\zeta'(\underline{x})$ assigned to at least one of these points could well be irrational.

2.3.2 D-optimality

Under the assumptions of normality of errors and linearity of the fitted models, a confidence ellipsoid for $\underline{\beta}$, of a given confidence coefficient has the form

$$\{\underline{\beta}: (\underline{\beta} - \hat{\underline{\beta}})'M(\zeta, \Sigma)(\underline{\beta} - \hat{\underline{\beta}}) \leq \text{constant}\}$$

where $\hat{\underline{\beta}}$ is the generalized least squares estimate of $\underline{\beta}$ (see Silvey [1980, p. 10] for the single response case). A more precise estimate for $\underline{\beta}$ is obtained by making this ellipsoid as small as possible. Since the volume of the ellipsoid is proportional to $|M(\zeta, \Sigma)|^{-1/2}$, this can be achieved by choosing a design measure which maximizes $|M(\zeta, \Sigma)|$. Such a design measure is said to be D-optimal and can be defined as follows.

Definition 2.2

Let H be the class of all design measures on χ. Then ζ^* is called D-optimal with respect to Σ if

$$|M(\zeta^*, \Sigma)| = \sup_{\zeta \varepsilon H} |M(\zeta, \Sigma)| \quad . \tag{2.15}$$

2.4 Construction of D-optimal Designs When Σ is Known

Fedorov's algorithm to construct D-optimal designs (when Σ is known) is a sequential procedure, in the sense that design points are chosen one at a time. It is based on a certain equivalence theorem. In this section we include the sequential procedure along with the statements of the equivalence theorem (Theorem 2.1) and the theorem concerning convergence of the sequential procedure. The interested reader is referred to Fedorov (1972, Ch. 5) for a more detailed discussion.

Theorem 2.1 (Equivalence Theorem) (Fedorov, 1972, p. 212)

Let H be the class of all design measures on χ, and let

$$V(\underline{x}, \zeta, \Sigma) = \phi'(\underline{x})M^{-1}(\zeta, \Sigma)\phi(\underline{x}) \quad , \quad \zeta \in H \quad . \qquad (2.16)$$

The following conditions are equivalent.

1. The design measure ζ^* is D-optimal.

2. $\sup_{\underline{x}\in\chi} tr[\Sigma^{-1}V(\underline{x}, \zeta^*, \Sigma)] = \inf_{\zeta\in H} \sup_{\underline{x}\in\chi} tr[\Sigma^{-1}V(\underline{x}, \zeta, \Sigma)]$.

3. $\sup_{\underline{x}\in\chi} tr[\Sigma^{-1}V(\underline{x}, \zeta^*, \Sigma)] = p$.

2.4.1 The Sequential Procedure When Σ is Known (Fedorov, 1972, Ch. 5)

For convenience, throughout the rest of this chapter we will write ζ_N for ζ_{D_N}. As described by Fedorov we start with a nondegenerate design D_{N_0}, i.e., $M(\zeta_{N_0}, \Sigma)$ is nonsingular. By successive addition of points to D_{N_0} we will construct a sequence of designs D_{N_0+1}, D_{N_0+2}, ... which converges to a D-optimal design. The steps in this procedure are as follows.

1. Find $\underline{x}_{N_0+1}$ in χ such that

$$tr[\Sigma^{-1}V(\underline{x}_{N_0+1}, \zeta_{N_0}, \Sigma)] = \sup_{\underline{x}\in\chi} tr[\Sigma^{-1}V(\underline{x}, \zeta_{N_0}, \Sigma)] \quad .$$

2. Form the new design D_{N_0+1} by adding $\underline{x}_{N_0+1}$ to D_{N_0}.

3. Continue this procedure and obtain the nested sequence of designs D_{N_0} D_{N_0+1} ... where D_N is obtained from D_{N-1} by adding the point $\underline{x}_N$ in χ which satisfies

$$\mathrm{tr}[\Sigma^{-1}V(\underline{x}_N, \zeta_{N-1}, \Sigma)] = \sup_{\underline{x}\varepsilon\chi} \mathrm{tr}[\Sigma^{-1}V(\underline{x}, \zeta_{N-1}, \Sigma)] \quad . \quad (2.17)$$

The stopping point for this procedure is reached when we find $\underline{x}_{N'}$ in χ which satisfies

$$\mathrm{tr}[\Sigma^{-1}V(\underline{x}_{N'}, \zeta_{N'-1}, \Sigma)] - p < \delta$$

where δ is a small positive number chosen a priori.

Theorem 2.2 (Fedorov, 1972, p.216)

The iterative procedure described above converges to a D-optimal design, i.e.,

$$\lim_{N\to\infty}|M(\zeta_N, \Sigma)| = |M(\zeta^*, \Sigma)|$$

where ζ^* is a D-optimal design measure.

2.5 Construction of D-optimal Designs When Σ is Not Known

As was clearly seen the moment matrix $M(\zeta, \Sigma)$ does depend on Σ. For this reason the above procedure cannot be applied unless Σ is known. If Σ is unknown, a consistent estimator of Σ can be used to

construct a sequence of random design measures which converges in probability to a D-optimal design measure with respect to Σ. (For definition of convergence in probability see Appendix A). Such a consistent estimator (see Appendix C for proof) was proposed by Zellner (1962) and is given by $\hat{\Sigma}_N = [\hat{\sigma}_{ij}^N]$ where

$$N\hat{\sigma}_{ij}^N = (\underline{Y}_i - F_i\underline{\tilde{\beta}}_i)'(\underline{Y}_j - F_j\underline{\tilde{\beta}}_j) \quad , \quad i,j = 1, 2, \ldots, r, \tag{2.18}$$

and $\underline{\tilde{\beta}}_i = (F_i'F_i)^{-1}F_i'\underline{Y}_i$ is the usual single equation least squares estimator, $i = 1, 2, \ldots, r$. Now suppose

$$A = (\text{diag } \Sigma^{-1})^{-1/2}\, \Sigma^{-1}(\text{diag } \Sigma^{-1})^{-1/2} \tag{2.19}$$

where diag(P) denotes the diagonal matrix, whose diagonal elements are equal to the diagonal elements of the square matrix P. Then the following theorem proves that a D-optimal design with respect to Σ is equivalent to a D-optimal design with respect to A.

Theorem 2.3

$$|M(\zeta, \Sigma)| = \prod_{i=1}^{r} (\sigma^{ii})^{p_i} |M(\zeta, A^{-1})|, \; \zeta \in H$$

where $\Sigma^{-1} = [\sigma^{ij}]$ and p_i is the number of parameters in the model $\underline{Y}_i = F_i\underline{\beta}_i + \underline{\varepsilon}_i$. The proof is given in Appendix D.

It can be seen from the expression on the right hand side that only $|M(\zeta, A^{-1})|$ depends on the design measure ζ. Hence a design measure which maximizes $|M(\zeta, A^{-1})|$ will essentially maximize $|M(\zeta, \Sigma)|$. It can be also shown that the results stated in Section 2.4 hold true when Σ^{-1} is replaced by A and thus an asymptotically D-optimal design measure with respect to A can be obtained using the sequential procedure described in Subsection 2.4.1. However, when Σ is not known, A is also not known and so we modify this procedure so that A is replaced by

$$\hat{A}_N = (\text{diag } \hat{\Sigma}_N^{-1})^{-1/2} \, \hat{\Sigma}_N^{-1} \, (\text{diag } \hat{\Sigma}_N^{-1})^{-1/2} \tag{2.20}$$

which will be updated every time a new design point is chosen. This procedure is described in Section 2.6.

The construction of a D-optimal design with respect to A (instead of Σ^{-1}) is more desirable for two reasons.

1. Since the elements of A lie between -1 and 1, $\hat{A}_N$ reaches stability much faster than $\hat{\Sigma}_N^{-1}$ thereby giving rise to a rapidly converging point in the sequential procedure described below.
2. Due to the fact that the diagonal elements of A are all equal to 1 the number of elements of A to be estimated is reduced by r.

2.5.1 The Sequential Procedure When Σ is Not Known

1. Start with an initial nondegenerate design D_{N_0} with N_0 points and compute $\hat{\Sigma}_{N_0}$ and $\hat{A}_{N_0}$ (based on observations on all r

responses measured at all the points of D_{N_0}) using (2.18) and (2.20).

2. Construct $\hat{D}_{N_0+1}$ by adding $\hat{\underline{x}}_{N_0+1}$ to D_{N_0} where $\hat{\underline{x}}_{N_0+1}$ satisfies

$$\mathrm{tr}[\hat{A}_{N_0} V(\hat{\underline{x}}_{N_0+1}, \zeta_{N_0}, \hat{A}^{-1}_{N_0}) \approx = \sup_{\underline{x}\varepsilon X} \mathrm{tr}[\hat{A}_{N_0} V(\underline{x}, \zeta_{N_0}, \hat{A}^{-1}_{N_0}) \approx \quad ,$$

and obtain $\hat{\zeta}_{N_0+1}$ by assigning mass $1/(N_0+1)$ to each design point in $\hat{D}_{N_0+1}$.

3. Once $\hat{D}_{N_0+N}$, $N > 1$, is obtained, compute $\hat{\Sigma}_{N_0+N}$ and $\hat{A}_{N_0+N}$ (based on the observations on all r responses measured at all the points of $\hat{D}_{N_0+N}$) using (2.18) and (2.20).

4. Construct $\hat{D}_{N_0+N+1}$ by adding $\hat{\underline{x}}_{N_0+N+1}$ to $\hat{D}_{N_0+N}$ which satisfies

$$\mathrm{tr}[\hat{A}_{N_0+N} V(\hat{\underline{x}}_{N_0+N+1}, \hat{\zeta}_{N_0+N}, \hat{A}^{-1}_{N_0+N}) \approx$$

$$= \sup_{\underline{x}\varepsilon X} \mathrm{tr}[\hat{A}_{N_0+N} V(\underline{x}, \hat{\zeta}_{N_0+N}, \hat{A}^{-1}_{N_0+N}) \approx$$

and obtain $\hat{\zeta}_{N_0+N+1}$, by assigning mass $1/(N_0+N+1)$ to all the design points in $\hat{D}_{N_0+N+1}$.

The stopping rule for this procedure is

$$\sup_{\underline{x}\varepsilon X} \mathrm{tr}[\hat{A}_{N'} V(\underline{x}, \hat{\zeta}_{N'}, \hat{A}^{-1}_{N'}) \approx - p < \delta$$

where δ is chosen beforehand and N' is some positive integer $> N_0$.

We now consider convergence of the procedure given above. Let **A** be the set of all $r \times r$ symmetric matrices $A = [a_{ij}]$ such that $a_{ii} = 1$, $i = 1, 2, \ldots, r$ and $-1 < a_{ij} < 1$, $1 \leq i < j \leq r$. Define $\underline{b}(A) = (b_1, b_2, \ldots, b_{r'})'$ where $r' = r(r-1)/2$, to be the vector which consists of the elements of A above its diagonal taken in order from left to right for each row starting with the first. We call $\underline{b}(A)$ the r'-dimensional vector associated with A. Note that since $A \epsilon \mathbf{A}$ is symmetric and all its diagonal elements are equal to 1, A can be completely described by the r' elements a_{ij}, $1 \leq i < j \leq r$. It is clear that A and $\{\hat{A}_N, N \geq 1\}$ defined in (2.19) and (2.20), respectively, belong to **A**. Also it is shown in Appendix E that $\hat{A}_N$ is a consistent estimator of A. Equivalently, if $\hat{e}_N$ is the Euclidean distance between $\underline{b}(\hat{A}_N)$ and $\underline{b}(A)$, then $\hat{e}_N$ converges in probability to 0.

It is conjectured that the convergence in probability to 0 of $\hat{e}_N$ implies that the above sequential procedure converges. However, we are able to prove convergence of the sequential procedure only by assuming a stronger condition that $\sum_{u=1}^{N} \hat{e}_u$ converges in probability to some random variable e. This can be formally stated as follows.

<u>Theorem 2.4</u>

Suppose that $\sum_{u=1}^{N} \hat{e}_u$ converges in probability to some random variable e. Then for a given $\delta > 0$, there exists an integer $N > 0$ such that

$$\sup_{\underline{x} \epsilon \chi} tr[\hat{A}_N V(\underline{x}, \hat{\zeta}_N, \hat{A}^{-1}_N)] - p < \delta$$

with probability 1 (see Appendix A for definition of convergence with probability 1).

The proof is given in Appendix F.

2.6 A Numerical Example

We now present an example to illustrate the procedure described in Section 2.6. We consider an experiment with two responses which depend on three controllable variables. The experimental region $\chi = \{\underline{x} = (x_1, x_2, x_3): |x_i| \leq 1.73,\ i = 1, 2, 3\}$ and the fitted models are

$$y_1 = \beta_{10} + \beta_{11}x_1 + \beta_{12}x_2 + \beta_{13}x_3 + \beta_{112}x_1x_2 + \beta_{113}x_1x_3 + \beta_{111}x_1^2 + \beta_{133}x_3^2 + \varepsilon_1 \tag{2.21}$$

$$y_2 = \beta_{20} + \beta_{21}x_1 + \beta_{22}x_2 + \beta_{212}x_1x_2 + \beta_{211}x_1^2 + \beta_{222}x_2^2 + \varepsilon_2 .$$

Thus $\underline{\beta}_1$ and $\underline{\beta}_2$ are of dimensions $p_1 = 8$ and $p_2 = 6$ and the total number of unknown parameters is $p = 14$. The matrices F_1, F_2 consist of eight and six columns, respectively, and the matrix F_{D_N} is of the form

$$F_{D_N} = \begin{bmatrix} F_1 & 0 \\ & \\ 0 & F_2 \end{bmatrix} .$$

The error vector $\varepsilon = (\underline{\varepsilon}_1', \underline{\varepsilon}_2')'$ is assumed to be normally distributed with zero mean vector and variance-covariance matrix $\Sigma \otimes I_N$ where

$$\Sigma = \begin{bmatrix} 2 & .4 \\ .4 & 1 \end{bmatrix}.$$

The matrix Σ will be used to generate the normal error values ε_1 and ε_2 with the stated variance-covariance matrix but will not be used in the sequential generation of the D-optimal design. Also

$$\Sigma^{-1} = \begin{bmatrix} .543 & -.217 \\ -.217 & 1.086 \end{bmatrix} \quad \text{and} \quad A = \begin{bmatrix} 1 & -.28 \\ -.28 & 1 \end{bmatrix}.$$

The error vector $\underline{\varepsilon}$ was generated by means of a SAS procedure which produces normally distributed random vectors. The initial design consisted of $N_0 = 15$ points and is given in Table 2.1. At the Nth iteration $\hat{\sigma}_{ij}^{N_0+N}$, $1 \le i, j \le 2$ was calculated by

$$(N_0+N)\hat{\sigma}_{ij}^{N_0+N} = \underline{\varepsilon}_i'(I - R_i)'(I - R_j)\underline{\varepsilon}_j \tag{2.22}$$

where $R_i = F_i(F_i'F_i)^{-1} F_i'$, $i = 1, 2$, and $\underline{\varepsilon}_k$, F_k, $k = 1, 2$, consists of (N_0+N) rows. Observe that in situations where response values are available, $\hat{\sigma}_{ij}^{N_0+N}$ should be calculated using (2.18). Using (2.20) $\hat{a}_{12}^{N_0+N}$ was then calculated and these values are given in Table 2.2. The maximization at each iteration was carried out using a computer program

Table 2.1
The Initial Design

x_1	x_2	x_3
1	1	-1
1	1	1
1	-1	-1
1	-1	1
-1	1	-1
-1	1	1
-1	-1	-1
-1	-1	1
1.68	0	0
0	1.68	0
0	0	0
0	0	0
0	0	0
0	0	0
0	0	0

Table 2.2

The Estimates for a_{12} and $\{\sigma^{ij},\ i,j=1,\ 2\}$

N_0+N-1	$-\hat{a}_{12}^{N_0+N-1}$	$\hat{\sigma}^{11}_{N_0+N-1}$	$\hat{\sigma}^{22}_{N_0+N-1}$	$-\hat{\sigma}^{12}_{N_0+N-1}$
15	.302	.752	.958	.256
16	.334	.722	1.03	.280
17	.220	.680	1.06	.186
18	.214	.648	1.01	.172
19	.024	.653	1.03	.018
20	.278	.657	1.36	.266
21	.214	.660	1.35	.202
22	.256	.670	1.30	.238
23	.218	.660	1.33	.204
24	.080	.703	1.32	.076
25	.090	.678	1.29	.084
26	.114	.652	1.30	.104
27	.212	.660	1.28	.194
28	.298	.638	1.29	.270
29	.292	.616	1.28	.260
30	.284	.596	1.33	.252
31	.298	.584	1.40	.270
32	.304	.623	1.55	.300
33	.284	.699	1.54	.294
34	.332	.619	1.89	.360
35	.318	.632	1.84	.344
36	.320	.621	1.85	.342
37	.332	.605	1.83	.350

Table 2.3
The Augmented Design Points for a D-Optimal Design

N_0+N	$tr\hat{A}_{N_0+N-1}V(\underline{x},\hat{\zeta}_{N_0+N-1},\hat{A}_{N_0+N-1})$	x_1	x_2	x_3
16	105.00	1.724	1.724	-1.728
17	101.47	1.729	1.727	-1.703
18	88.74	1.728	-1.729	-1.720
19	84.29	-1.728	-1.725	-1.680
20	50.73	1.729	1.729	1.729
21	52.00	-1.725	-1.723	1.715
22	42.42	-1.730	1.721	1.729
23	48.80	1.730	-1.729	1.729
24	22.95	-1.730	1.730	.026
25	23.95	1.708	1.660	1.692
26	23.35	1.730	-1.730	-.045
27	25.19	-1.729	-1.730	-1.728
28	20.55	-1.730	-.096	1.730
29	23.07	1.729	1.724	-1.729
30	21.34	1.730	-1.730	.153
31	20.58	-.154	1.730	-1.730
32	21.14	1.690	-1.702	1.658
33	20.48	1.727	-1.710	-1.726
34	20.76	-.101	-1.730	1.730
35	19.92	-1.723	1.601	1.695
36	20.00	1.729	1.729	1.722
37	16.56	1.729	-1.703	1.698

called SEARCH written by Conlon (1979). The program is based on the controlled random search procedure introduced by Price (1977). This procedure uses a random search to locate an optimal point from among a collection of points, the number of which is predetermined by the user. The augmented design points are given in Table 2.3. Observe that the procedure has succeeded in achieving a substantial reduction in the 'tr' value, (105 to 16.56). Also the difference between the final value of 16.56 and the anticipated value of 14 is relatively small. In addition it can be seen that the values of $\hat{a}_{12}$ have stabilized. This demonstrates the effectiveness of our procedure in constructing D-optimal designs when Σ is not known.

In Table 2.3 the last entry indicates a 'tr' value of 16.56 at the 22nd iteration. When further iterations were carried out no further drop was noted in the 'tr' value. This may be due to the fact that the number of points used for the random search may have been insufficient to locate the true optimal point.

CHAPTER THREE
DESIGN CRITERIA TO INCREASE THE POWER OF LACK OF FIT TEST

3.1 Introduction

In the past, considerable attention has been focussed on the topic of model inadequacy in linear regression. Its importance is described in Box and Draper (1959). Once a single response experiment is performed assuming a certain model is true, a test to detect lack of fit of the fitted model can be contemplated (see Draper and Smith, 1981, Ch. 2). Authors such as Aitkinson (1972), Aitkinson and Federov (1975) and Jones and Mitchell (1978) have come up with design procedures to improve the detection of model inadequacy in single response experiments.

Detection of model inadequacy is an important issue in the multiresponse case as it is in the single response case. Khuri (1983) developed a test for lack of fit for a linear multiresponse model. As mentioned by Khuri (1983), the inadequacy of the fitted model can seriously affect the estimation procedure of regression coefficients developed by Box and Draper (1965). A detailed discussion of this procedure and the problems associated with it is given in Box and Draper (1965), Box et al. (1973) and McLean et al. (1979). Khuri (1983) also

states that the multiresponse lack of fit test provides a comprehensive assessment of the adequacy of the multiresponse model. It can detect lack of fit of the overall model even in situations where the usual single response lack of fit test performed on each individual response separately fails to reveal any inadequacies in the corresponding model. This fact was well illustrated by Khuri, by performing the multiresponse lack of fit test on data obtained from a study made by Kizer et al. (1978) on the behavior of a fluidized bed reactor for the catalytic oxidation of benzene to mallic anhydride. Khuri (1983) also gave a procedure to determine which responses are responsible for lack of fit when the multiresponse test for lack of fit is significant.

The purpose of this chapter is to develop design criteria to increase the power of the multiresponse test of lack of fit mentioned above. The model assumptions and the notation in this chapter follow Khuri (1983) and are given below for completeness.

3.2 Model and Notation

Let N be the total number of experimental runs and r be the number of responses. We assume that each response is dependent on all or some of the k controllable variables denoted x_1, x_2, ..., x_k. The dependencies of the responses on the controllable variables are expressible in terms of polynomials of possibly different degrees. The fitted i^{th} response model can be represented as

$$E_a(\underline{Y}_i) = X_i\underline{\beta}_i, \quad i = 1, 2, \ldots, r \tag{3.1}$$

where $\underline{Y}_i$ is an Nx1 vector of observations on the i^{th} response, $E_a(\underline{y}_i)$ denotes the expected value of $\underline{Y}_i$ under the fitted model (the subscript 'a' denotes assumed), X_i is an Nxp_i matrix of rank p_i of known functions of the settings of the controllable variables, and $\underline{\beta}_i$ is a p_ix1 vector of unknown constant parameters (i = 1, 2, ..., r).

The model for the true i^{th} response mean (i = 1, 2, ..., r) is assumed to be of the same form as the fitted model but possibly contains terms in addition to those in the fitted model. It can be written as follows.

$$E_t(\underline{Y}_i) = X_i\underline{\beta}_i + Z_i\underline{\gamma}_i, \quad i = 1, 2, \ldots, r \tag{3.2}$$

where $E_t(\underline{Y}_i)$ denotes the expected value of $\underline{Y}_i$ under the true model, Z_i is an Nxq_i matrix of known functions of the settings of the controllable variables, and $\underline{\gamma}_i$ is a vector of unknown constant parameters. If the fitted model (3.1) is correct, then $\underline{\gamma}_i$ will be equal to the zero vector.

The equations given in (3.1) and (3.2) can be written to represent the overall multiresponse model in the following manner.

$$E_a(Y) = XB \tag{3.3}$$

$$E_t(Y) = XB + Z\Gamma \tag{3.4}$$

where $Y = [\underline{Y}_1 : \underline{Y}_2 : \ldots : \underline{Y}_r]$, $X = [X_1 : X_2 : \ldots : X_r]$, $Z = [Z_1 : Z_2 : \ldots : Z_r]$, $B = \mathrm{Diag}[\underline{\beta}_1, \underline{\beta}_2, \ldots, \underline{\beta}_r]$, and $\Gamma = \mathrm{Diag}[\underline{\gamma}_1, \underline{\gamma}_2, \ldots, \underline{\gamma}_r]$. The matrices Y, X, Z, and Γ are of orders Nxr, Nxp, Nxq, and qxr, respectively, where $p = \sum_{i=1}^{r} p_i < N$, $q = \sum_{i=1}^{r} q_i$, and X is of rank $\rho(\leq p)$. The rows of Y are independent observations from multivariate normal populations with a common nonsingular variance-covariance matrix Σ of order rxr. Under the true model, Y has a mean given by (3.4), and is referred to as the normal data matrix with mean $XB + Z\Gamma$ and variance-covariance matrix $I_N \otimes \Sigma$. This is written symbolically as

$$Y \sim N(XB + Z\Gamma,\ I_N \otimes \Sigma) \quad .$$

In the development of the lack of fit test it is assumed that replicated observations are available on all r responses. Without loss of generality it will be assumed that such replicated observations are obtained at each of the first n design points, where $1 \leq n \leq N$. The number of repeated observations at the ith design point is denoted by ν_i ($\nu_i \geq 2$, $i = 1, 2, \ldots, n$). Let $\nu = \sum_{i=1}^{n} \nu_i$ ($\nu \leq N$). The matrices Y, X, and Z in (3.3) and (3.4) are then partitioned as

$$Y = \begin{bmatrix} Y^{(1)} \\ \\ Y^{(2)} \end{bmatrix} \quad , \quad X = \begin{bmatrix} X^{(1)} \\ \\ X^{(2)} \end{bmatrix} \quad , \quad Z = \begin{bmatrix} Z^{(1)} \\ \\ Z^{(2)} \end{bmatrix}$$

where $Y^{(1)}$, $X^{(1)}$, and $Z^{(1)}$ are matrices of orders $\nu x r$, $\nu x p$, and $\nu x q$, respectively, and are associated with the replicated portion of the design, whereas the matrices $Y^{(2)}$, $X^{(2)}$, and $Z^{(2)}$ are associated with the nonreplicated portion.

3.3 Development of Khuri's Lack of Fit Test

Khuri (1983) developed a multivariate lack of fit test by first transforming the normal data matrix Y into a single response normal data vector and applying the usual single response test for lack of fit described in Draper and Smith (1981, Ch. 2). The null hypothesis tested by the multivariate lack of fit test is

$$H_0\colon [I_N - X_0(X_0'X_0)^{-1}X_0']Z\Gamma = 0 \tag{3.5}$$

where X_0 is a $N x \rho$ matrix of rank ρ whose columns form a basis for the columns of X, which is of rank ρ.

The null hypothesis H_0 is the hypothesis of zero lack of fit (or adequacy of fit) of the model (3.3) when in reality the true model is given by (3.4). The test statistics to test this null hypothesis as developed by Khuri (1983) is explained below. Let

$$K = \mathrm{Diag}(K_1, K_2, \ldots, K_n, 0)$$

where K is a block diagonal matrix of order NxN, 0 is a zero matrix of order $(N-\nu) \times (N-\nu)$, and

$$K_i = I_{\nu_i} - (1/\nu_i)J_{\nu_i} \quad , \quad i = 1, 2, \ldots, n \quad . \tag{3.7}$$

Here I_{ν_i} is the identity matrix of order $\nu_i x \nu_i$ and J_{ν_i} is the matrix of ones of order $\nu_i x \nu_i$ $(i = 1, 2, \ldots, n)$.

Define the matrices G_1 and G_2 as

$$G_1 = Y'[I_N - X_0(X_0'X_0)^{-1}X_0' - K]Y \tag{3.8}$$

and

$$G_2 = Y'KY \quad . \tag{3.9}$$

An appropriate test statistic to test H_0 given in (3.5) is the maximum eigenvalue of $G_1G_2^{-1}$, denoted by $e_{max}(G_1G_2^{-1})$, and the test procedure is reject H_0 at the α-level of significance if

$$e_{max}(G_1G_2^{-1}) > \lambda_\alpha \tag{3.10}$$

where λ_α is the upper 100α % point of the null distribution of $e_{max}(G_1G_2^{-1})$ (i.e., the distribution of $e_{max}(G_1G_2^{-1})$ under H_0). This multivariate test is known as Roy's largest root test. Three other multivariate test statistics are also available to test H_0. They are (1) Wilks' likelihood ratio, $W = |G_2|/|G_1 + G_2|$, (2) Pillai's trace,

$tr[G_1(G_1+G_2)^{-1}]$, and (3) Hotelling-Lawley's trace, $tr(G_1G_2^{-1})$ where '$|\ |$' and 'tr' respectively denote the determinant and the trace of a matrix.

3.4 Development of Design Criteria

It is shown by Khuri (1983) that G_2 has the central Wishart distribution with $\nu_{PE} = \sum_{i=1}^{n} (\nu_i - 1)$ degrees of freedom; also that G_1 is independent of G_2 and has the noncentral Wishart distribution with $\nu_{LF} = (N-\rho-\nu_{PE})$ degrees of freedom and a noncentrality parameter matrix given by

$$\Omega = \Sigma^{-1}\Gamma'Z'[I_N - X_0(X_0'X_0)^{-1}X_0']Z\Gamma \quad . \tag{3.11}$$

It is known that the power of the lack of fit test based upon any of these four test statistics mentioned above is a montone increasing function of the eigenvalues of Ω (see Roy et al., 1971, p. 68). Therefore, the power of this test can be increased by increasing the eigenvalues of Ω. This can in turn be achieved by maximizing the trace of Ω. This is true since increasing the trace of Ω will increase at least one of its eigenvalues. However, the choice of the design which maximizes the trace of Ω depends on the matrices Σ and Γ which are unknown. Thus we are faced with the problem of finding an expression independent of Σ and Γ, whose maximization will result in an increase in the trace of Ω. This expression is found as follows.

From Theorem B.3 in Appendix B we know that

$$tr(\Omega) = tr\{\Sigma^{-1/2}\Gamma'Z'[I_N-X_0(X_0'X_0)^{-1}X_0']Z\Gamma\Sigma^{-1/2}\}$$

$$\geq e_{min}(\Sigma^{-1/2})tr\{\Gamma'Z'[I_N-X_0(X_0'X_0)^{-1}X_0']Z\Gamma\} \quad . \qquad (3.12)$$

Recall that $Z = [Z_1 : Z_2 : \ldots : Z_r]$ where the columns of Z_i, correspond to the extra terms in the true model for the i^{th} response ($i = 1, 2, \ldots, r$). It is easy to see that more than one column in Z can be assigned to the same term in the extra portion of the multiresponse model i.e., more than one model can have the same term. Let Z_0 be the matrix obtained from Z by deleting those columns which occur more than once so that each term in the true model corresponds to exactly one column in Z_0. For example let us consider a situation with two responses and let the columns of Z_1 and Z_2 correspond to terms $\{x_1^2, x_2^2\}$ and $\{x_1^2, x_2^2, x_1x_2\}$, respectively. Then Z consists of five columns with columns one and three corresponding to x_1^2, columns two and four corresponding to x_2^2 and column five corresponding to x_1x_2 so that Z_0 is obtained by deleting columns three and four.

In general, Z_i, $i = 1, 2, \ldots, r$, can be written as

$$Z_i = Z_0H_i \quad , \quad i = 1, 2, \ldots, r \quad . \qquad (3.13)$$

Here Z_0 is of order $N x \rho_1$ where ρ_1 is the number of columns in Z_0. The H_i matrices are of order $\rho_1 x q_i$ and their elements consist of zeros and ones. Now

$$tr\{\Gamma' Z' [I_N - X_0(X_0'X_0)^{-1}X_0']Z\Gamma\}$$

$$= tr \begin{bmatrix} \underline{\gamma}_1' & & & \\ & \underline{\gamma}_2' & & \\ & & \ddots & \\ & & & \underline{\gamma}_r' \end{bmatrix} \begin{bmatrix} H_1'Z_0' \\ H_2'Z_0' \\ \vdots \\ H_r'Z_0' \end{bmatrix} [I_N - X_0(X_0'X_0)^{-1}X_0'][Z_0H_1 : \dots : Z_0H_r] \begin{bmatrix} \underline{\gamma}_1 & & & \\ & \underline{\gamma}_2 & & \\ & & \ddots & \\ & & & \underline{\gamma}_r \end{bmatrix}$$

$$= tr \begin{bmatrix} \underline{\gamma}_1' H_1' \\ \underline{\gamma}_2 H_2' \\ \vdots \\ \underline{\gamma}_r' H_r \end{bmatrix} Z_0'[I_N - X_0(X_0X_0')^{-1}X_0']Z_0[H_1\underline{\gamma}_1 : H_2\underline{\gamma}_2 : \dots : H_r\underline{\gamma}_r]$$

$$= \sum_{i=1}^{r} \underline{\gamma}_i' H_i' A H_i \underline{\gamma}_i \quad \text{where} \quad A = Z_0'(I_N - X_0(X_0'X_0)^{-1}X_0')Z_0$$

$$= [\underline{\gamma}_1' H_1' : \underline{\gamma}_2' H_2' : \dots : \underline{\gamma}_r' H_r'](I_r \otimes A) \begin{bmatrix} H_1\underline{\gamma}_1 \\ H_2\underline{\gamma}_2 \\ \vdots \\ H_r\underline{\gamma}_r \end{bmatrix}$$

where the vector $[\underline{\gamma}_1' H_1' : \underline{\gamma}_2' H_2' : \ldots : \underline{\gamma}_r' H_r']$ contains the elements of γ_i, $i = 1, 2, \ldots, r$ and possibly some zeros. We want to move all the zero elements of $[\underline{\gamma}_1' H_1' : \underline{\gamma}_2' H_2' : \ldots : \underline{\gamma}_r' H_r']$ to the right and the nonzero ones to the left. This can be accomplished by writing

$$[\underline{\gamma}_1' H_1' : \underline{\gamma}_2' H_2' : \ldots : \underline{\gamma}_r' H_r'] = [\underline{\gamma}' : \underline{0}']C \tag{3.14}$$

where $\underline{\gamma}' = [\underline{\gamma}_1' : \underline{\gamma}_2' : \ldots : \underline{\gamma}_r']_{1xq}$, $\underline{0}$ is a $(r\rho_1 - q)x1$ vector of zeros, and C is an orthogonal matrix which has the function of reordering the elements of $[\underline{\gamma}_1' H_1' : \underline{\gamma}_2' H_2' : \ldots \underline{\gamma}_r' H_r']$. It is obtained by a suitable rearrangement of the columns of the identity matrix.

Therefore

$$\mathrm{tr}\{\Gamma' Z' [I_N - X_0 (X_0' X_0)^{-1} X_0'] Z \Gamma\}$$

$$= [\underline{\gamma}' : \underline{0}']\, C (I_r \otimes A) C' \begin{bmatrix} \underline{\gamma} \\ \underline{0} \end{bmatrix}$$

$$= \underline{\gamma}' [I_q : 0_{qx(r\rho_1 - q)}] C (I_r \otimes A) C' \begin{bmatrix} I_q \\ 0_{(r\rho_1 - q)xq} \end{bmatrix} \underline{\gamma}$$

$$= \underline{\gamma}' L (I_r \otimes A) L' \underline{\gamma} \quad \text{where} \quad L = [I_q : 0_{qx(r\rho_1 - q)}]C \quad . \tag{3.15}$$

Thus, inequality (3.12) can be written as

$$tr(\Omega) \geq e_{min}(\Sigma^{-1/2})\underline{\gamma}' \, L \, (I_r \otimes A) \, L'\underline{\gamma} \quad .$$

Now since $e_{min}(\Sigma^{-1/2})$ is a constant, the above inequality implies that maximization of the quantity

$$\Lambda = \underline{\gamma}' L(I_r \otimes A)L'\underline{\gamma}$$

will result in increasing the trace of Ω. However, the choice of design to maximize Λ depends on $\underline{\gamma}$, which is unknown. In order to overcome this problem we apply the maximin method which was proposed by Aitkinson and Fedorov (1975) and used by Jones and Mitchell (1978) in the single response case. The maximin method consists of choosing a design which maximizes Λ_1, the minimum of Λ over a specified region $\bar{\Lambda}$ in the $\underline{\gamma}$-space. The specification of this region $\bar{\Lambda}$ depends on a quantity τ which can be considered as a measure of the inadequacy of the fitted model. Suppose the u^{th} rows of X_i and Z_i, ($i = 1, 2, \ldots, r$; $u = 1, 2, \ldots, N$) in (3.2) can be represented by $\underline{f}_i'(\underline{x}_u)$ and $\underline{g}_i'(\underline{x}_u)$, respectively. Then the fitted and true response functions associated with (3.1) and (3.2) are $\underline{f}_i'(\underline{x})\beta_i$ and $\underline{f}_i'(\underline{x})\underline{\beta}_i + \underline{g}_i'(\underline{x})\gamma_i$, $i=1, 2, \ldots, r$ respectively. We express τ as

$$\tau = \underline{\gamma}' \, T \, \underline{\gamma} \quad , \quad \text{where } \gamma' = [\underline{\gamma}_1' : \underline{\gamma}_2' : \ldots : \underline{\gamma}_r'] \qquad (3.16)$$

and

$$T = \begin{bmatrix} T_1 & & & \\ & T_2 & & \\ & & \ddots & \\ & & & T_r \end{bmatrix}, \text{ with } T_i = \mu_{22}^i - \mu_{21}^i (\mu_{11}^i)^{-1}\mu_{12}^i,\ i = 1, 2, \ldots, r.$$

Here μ_{kl}^i, $k,l = 1, 2$, are the region moment matrices defined by $\mu_{22}^i = S \int_X \underline{g}_i(\underline{x})\underline{g}_i'(\underline{x})d\underline{x}$, $\mu_{21}^i = S \int_X \underline{g}_i(\underline{x})\underline{f}_i'(\underline{x})d\underline{x}$, $\mu_{11}^i = S \int_X \underline{f}_i(\underline{x})\underline{f}_i'(\underline{x})d\underline{x}$, and $\mu_{12}^i = S \int_X f_i(\underline{x})\underline{g}_i'(\underline{x})d\underline{x}$ where $S^{-1} = \int_X d\underline{x}$ and χ denotes the experimental region. This is the multiresponse extension of the expression for τ given by Jones and Mitchell (1978). Now since T_i, $i = 1, 2, \ldots, r$ is positive definite, T is also positive definite. Hence it is clear that τ will be equal to zero if and only if $\underline{\gamma}$ is equal to zero. This happens when the fitted model is true. So τ is a measure of the inadequacy of the fitted multiresponse model given in (3.3) when in reality the true model is (3.4). It is positive whenever the fitted model is inadequate and is zero otherwise.

3.4.1 Λ_1-optimality

Since τ will be positive if the fitted model is inadequate, then in such a case $\tau > \delta$ for some constant $\delta > 0$. We define $\bar{\Lambda} = \{\underline{\gamma}: \underline{\gamma}'T\underline{\gamma} > \delta\}$. The first design criterion is to maximize Λ_1 where

$$\Lambda_1 = \inf_{\underline{\gamma}\in\bar{\Lambda}} \underline{\gamma}'L(I_r \otimes A)L'\underline{\gamma} \quad .$$

Recall that $A = Z_0'[I_N - X_0(X_0'X_0)^{-1}X_0']Z_0$.

This is a multiresponse extension of the Λ_1 optimality proposed by Jones and Mitchell (1978) and the minimum value of the quadratic form $\underline{\gamma}'L(I_r \otimes A)L'\underline{\gamma}$ over $\bar{\Lambda}$ occurs on the boundary of $\bar{\Lambda}$ which is the contour $\tau = \delta$, since $\frac{\partial \Lambda_1}{\partial \underline{\gamma}} = \underline{0}$ only at the origin which is not an interior point of $\bar{\Lambda}$. Furthermore, as in Jones and Mitchell (1978), it can be shown that

$$\Lambda_1 = \delta e_{min}[T^{-1}L(I_r \otimes A)L'] \quad . \tag{3.17}$$

Since δ is a constant, a design which maximizes $e_{min}[T^{-1}L(I_r \otimes A)L']$ is called a Λ_1-optimal design. However, there are some situations in which $e_{min}[T^{-1}L(I_r \otimes A)L']$ is equal to zero for any choice of design. For instance this happens when the matrix $[X_0:Z_0]$ is not of full column rank since when $[X_0: Z_0]$ is less than full column rank A will be singular, and hence $e_{min}(A) = 0$. This implies that $e_{min}(I_r \otimes A) = 0$. Thus $I_r \otimes A$ is positive semidefinite, and $T^{-1/2}L(I_r \otimes A)L'(T^{-1/2})'$ is also positive semidefinite, which implies that $e_{min}[T^{-1}L(I_r \otimes A)L'] = 0$. Thus, it is clear that Λ_1-optimal designs can be obtained only under some special conditions. This leads us to propose a second design criterion which can be applied in more general situations.

3.4.2 Λ_2-optimality

Our second design criterion is to maximize Λ_2, the average of Λ (instead of the infimum of Λ) over the contour $\tau = \delta$, i.e., we propose to select a design which maximizes

$$\Lambda_2 = \int_{\bar{\Lambda}_0} \underline{\gamma}'L(I_r \otimes A)L'\underline{\gamma}dG / \int_{\bar{\Lambda}_0} dG \tag{3.18}$$

where dG is the differential of the area on the surface of the ellipsoid $\bar{\Lambda}_0 = \{\underline{\gamma}: \underline{\gamma}'T\underline{\gamma} = \delta\}$. Jones and Mitchell (1978) showed that the following equality holds,

$$\Lambda_2 = q^{-1}\delta\Lambda_2' \quad \text{where} \quad \Lambda_2' = tr[T^{-1}L(I_r \otimes A)L'] \quad . \tag{3.19}$$

A design which maximizes Λ_2 (respectively Λ_2') is called a Λ_2-optimal (respectively Λ_2'-optimal) design. Since q and δ are constants it is clear that Λ_2-optimal designs and Λ_2'-optimal designs are equivalent.

If the number of design points N is fixed beforehand a Λ_2-optimal design could be obtained by maximizing Λ_2' with respect to the Nk design settings (coordinates of the N design points). However, this may lead to computational difficulties especially for large values of N or k. A sequential procedure by which design points can be chosen one at a time is therefore quite desirable. Silvey (1980) came up with a sequential procedure to obtain a ϕ-optimal design, if ϕ is a real valued function of the moment matrix associated with a linear single response model, provided ϕ satisfied certain properties. In our case it can be shown that Λ_2' is a function of the moment matrix associated with a linear single response model, and also that Λ_2' satisifies the required properties. As shown in the next section, Silvey's procedure can thus be employed to obtain a Λ_2'-optimal design sequentially.

3.5 Sequential Generation of Λ_2'- Designs

Let us consider the 'artificial' single response model

$$\underline{y}^a = [X_0 \quad Z_0]\,\underline{\theta} + \underline{\varepsilon} \tag{3.20}$$

where $\underline{y}^a$ is a Nx1 vector of observed response values, X_0 and Z_0 are as given before, $\underline{\theta}$ is the parameter vector associated with the model, and $\underline{\varepsilon}$ is the Nx1 random error vector with $E(\underline{\varepsilon}) = \underline{0}$ and $Var(\underline{\varepsilon}) = I$. Let $\underline{x} = (x_1, x_2, \ldots, x_k)'$, where $\underline{x}$ is a design point of the experimental region, χ, $\underline{a}'(\underline{x}_u)$ and $\underline{b}'(\underline{x}_u)$ be the row vectors which represent the u^{th} rows of X_0 and Z_0 respectively, and $\mathbf{H}$ be the set of all design measures on χ (see section 2.3.1 for the definition of a design measure). If $M(\zeta)$ is the moment matrix associated with (3.20), then $M(\zeta)$ is of order mxm where $m = \rho + \rho_1$, and using (2.14) with $r = 1$, $\Sigma = 1$ and $\phi(\underline{x}) = \underline{h}(\underline{x})$, where $\underline{h}'(\underline{x}) = [a'(\underline{x}) : \underline{b}'(\underline{x})]$ we have

$$M(\zeta) = \sum_{u=1}^{s} \lambda_u \underline{h}(\underline{x}_u)\underline{h}'(x_u) \quad . \tag{3.21}$$

Here, $1 \leq s \leq m'$, $0 \leq \lambda_u \leq 1$ with $\sum_{u=1}^{s} \lambda_u = 1$, and $m' = (m(m+1)/2)+1$. Equation (3.21) can be rewritten as

$$M(\zeta) = \sum_{u=1}^{m'} \lambda_u \underline{h}(\underline{x}_u)\underline{h}'(\underline{x}_u) \tag{3.22}$$

where $\underline{\lambda}_{m'} = (\lambda_1, \lambda_2, \ldots, \lambda_{n'})'$ belongs to $\mathbf{U}$ and $\mathbf{U} = \{\underline{\lambda}: 0 \leq \lambda_u \leq 1$ and $\sum_{u=1}^{m'} \lambda_u = 1\}$. (Observe that $\lambda_u = 0$ for $s < u \leq m'$).

Alternatively,

$$M(\zeta) = \begin{bmatrix} M_{XX}(\zeta) & M_{XZ}(\zeta) \\ \\ M_{ZX}(\zeta) & M_{ZZ}(\zeta) \end{bmatrix} \tag{3.23}$$

where $M_{XX}(\zeta) = \sum_{u=1}^{m'} \lambda_u \underline{a}(\underline{x}_u)\underline{a}'(x_u)$, $M_{XZ}(\zeta) = \sum_{u=1}^{m'} \lambda_u \underline{a}(x_u)\underline{b}'(\underline{x}_u)$, $M_{XZ}(\zeta) = \sum_{u=1}^{m'} \lambda_u \underline{b}(\underline{x}_u)\underline{a}'(x_u)$, and $M_{ZZ}(\zeta) = \sum_{u=1}^{m'} \lambda_u \underline{b}(x_u)\underline{b}'(\underline{x}_u)$. Let $\mathbf{M} = \{M(\zeta): \zeta \in H\}$. Then for $\mathbf{M}(\zeta) \in \mathbf{M}$ we can associate a point $\underline{\omega} = (\lambda_1, \lambda_2, \ldots, \lambda_{m'}, \underline{x}_1, \underline{x}_2, \ldots, \underline{x}_{m'})'$ in the closed and bounded subspace $\mathbf{U} \times \chi^{m'}$. Let D_N denote the set of N design points for fitting model (3.20). Then we can form a design measure ζ_{D_N} associated with D_N, by attaching probability 1/N to each design point, and using (2.13) with $r=1$, $\Sigma=1$, and $F_{D_N} = [X_0: Z_0]$, we have

$$M(\zeta_{D_N}) = \begin{bmatrix} M_{XX}(\zeta_{D_N}) & M_{XZ}(\zeta_{D_N}) \\ \\ M_{ZX}(\zeta_{D_N}) & M_{ZZ}(\zeta_{D_N}) \end{bmatrix} \tag{3.24}$$

where $M_{XX}(\zeta_{D_N}) = X_0'X_0/N$, $M_{XZ}(\zeta_{D_N}) = X_0'Z_0/N$, $M_{ZX}(\zeta_{D_N}) = Z_0'X_0/N$, and $M_{ZZ}(\zeta_{D_N}) = Z_0'Z_0/N$.

Recall that $\Lambda_2' = tr[T^{-1}L(I_r \otimes A)L')]$, where $A=Z_0'(I_N-X_0(X_0'X_0)^{-1}X_0')Z_0$. Now since $A = N\{M_{ZZ}(\zeta_{D_N}) - M_{ZX}(\zeta_{D_N})M_{XX}^{-1}(\zeta_{D_N})M_{XZ}(\zeta_{D_N})\}$, Λ_2' associated

with an N-point design is a function of the moment matrix $M(\zeta_{D_N})$. Thus the generalization of the Λ_2' function for any $M(\zeta) \in \mathbf{M}$ can be written as

$$\Lambda_2'[M(\zeta)] = \mathrm{tr}[T^{-1}L\{I_r \otimes (M_{ZZ}(\zeta) - M_{ZX}(\zeta)M_{XX}^{-1}(\zeta)M_{XZ}(\zeta)\}L'] \quad . \tag{3.26}$$

Thus Λ_2' is a function of the moment matrix associated with a linear single response model.

We shall now describe Silvey's sequential procedure. The relevant definitions are given first, and the notation used is consistent with model (3.20). For convenience, we shall write ζ_{D_N} for ζ_N.

Let ϕ be a real valued function bounded above on $\mathbf{M}$ but not necessarily bounded below, i.e., $\phi[M(\zeta)] = -\infty$ for some design measures $\zeta \in \mathbf{H}$. Let $\mathbf{M}' = \{M(\zeta): \phi[M(\zeta)] > -\infty\}$.

<u>Definition 3.1</u>

ϕ is said to be concave on $\mathbf{M}$, if for $0 < \alpha < 1$ and M^a, $M^b \in \mathbf{M}$ $\phi(M) \geq \alpha\phi(M^a) + (1 - \alpha)\phi(M^b)$, where $M = \alpha M^a + (1 - \alpha)M^b$.

<u>Definition 3.2</u>

Let $M(\zeta) \in \mathbf{M}'$. Then ϕ is said to be differentiable at $M(\zeta)$, if ϕ is partially differentiable with respect to each element of $\underline{\omega}$, where $\underline{\omega}$ is the point in $\mathbf{U} \times X^{m'}$ which is associated with $M(\zeta)$.

Definition 3.3

A design measure ζ^* is said to be ϕ-optimal if and only if

$$\phi[M(\zeta^*)] = \sup_{\zeta \varepsilon H} \phi[M(\zeta)] \quad .$$

Definition 3.4

Let M_1, M_2 ε **M**. Then the Frechet derivative of ϕ at M_1 in the direction of M_2 is:

$$F_\phi(M_1, M_2) = \lim_{\varepsilon \to 0^+} (1/\varepsilon)[\phi\{(1 - \varepsilon)M_1 + \varepsilon M_2\} - \phi(M_1)] \quad .$$

For the development of the theory which leads to Silvey's procedure mentioned earlier, it is necessary that ϕ be concave on **M** and differentiable on **M'** (Silvey, 1980, pp. 17-18). In addition, a ϕ-optimal measure should exist (Silvey, 1980, p. 22). The basic idea used in constructing the sequential procedure (Silvey, 1980, p. 29) is to choose the design point $\underline{x}_{N+1}$ such that

$$F_\phi\{M(\zeta_N), h(\underline{x}_{N+1})\underline{h}'(\underline{x}_{N+1})\} = \max_{\underline{x} \varepsilon X} F_\phi\{M(\zeta_N), h(x)\underline{h}'(\underline{x})\} \quad .$$

The stopping rule is $\max\limits_{x \varepsilon X} F_\phi\{M(\zeta_N), \underline{h}(x)\underline{h}'(\underline{x})\} = 0$, which is based on the following lemma (Silvey, 1980, p. 22): Let ϕ be concave on **M** and differentiable on **M'**. Suppose a ϕ-optimal measure exists. Then ζ^* is ϕ-optimal if and only if

$$\max_{\underline{x} \in X} F_\phi\{M(\zeta^*) \quad , \quad \underline{h}(\underline{x})\underline{h}'(\underline{x})\} = 0 \quad .$$

The main steps of the procedure are given below.

3.5.1. The Sequential Procedure to Obtain a ϕ-Optimal Design

1. Start with an initial design D_{N_0} such that $M(\zeta_{N_0}) \in \mathbf{M}'$.

2. Obtain the design point $\underline{x}_{N_0+1}$ at which

$$\max_{\underline{x} \in X} F_\phi[M(\zeta_{N_0}), \underline{h}(\underline{x})\underline{h}'(\underline{x})] \text{ is attained.}$$

3. Obtain D_{N_0+1} (hence ζ_{N_0+1}) by augmenting D_{N_0} with $\underline{x}_{N_0+1}$. Recall that ζ_{N_0+1} is the design measure obtained by assigning probability $1/(N_0+1)$ to each design point in D_{N_0+1}.

4. Repeat this process to find $\underline{x}_{N_0+2}$, $\underline{x}_{N_0+3}$, ... by replacing ζ_{N_0}, by ζ_{N_0+1}, ζ_{N_0+2}, ... until

$$\max_{\underline{x} \in X} F_\phi[M(\zeta_N), \underline{h}(x)\underline{h}'(\underline{x})] < \delta$$

for some N and δ, where δ is a small positive number chosen a priori.

The design measure ζ_{N+1}, defined iteratively in the above sequential procedure, gives equal weights $1/(N+1)$ to each of its support points and hence satisfies the recursive formula

$$\zeta_{N+1} = (1 - \alpha_N)\zeta_N + \alpha_N \zeta(\underline{x}_{N+1})$$

where $\alpha_N = 1/(N+1)$ and $\zeta(\underline{x}_{N+1})$ denotes the design measure which assigns weight 1 to the point $\underline{x}_{N+1}$. This formula is interpreted as follows. ζ_{N+1} assigns probability α_N to the design point $\underline{x}_{N+1}$ and a cumulative probability $(1 - \alpha_N)$ to $D_N = \{x_1, \underline{x}_2, \ldots, \underline{x}_N\}$ (i.e., the probabilities assigned to $\underline{x}_i$, i=1, 2, ..., N sum to $(1 - \alpha_N)$). Also $\lim_{N\to\infty} \alpha_N = 0$ and $\sum_{u=1}^{N} \alpha_u$ is divergent and for such $\{\alpha_N\}$ the above sequential procedure converges. This is formally stated in the following theorem.

Theorem 3.2

The procedure described in section 3.5.1 converges, i.e., given $\delta > 0$,

$$\max_{\underline{x}\in\chi} F_\phi[M(\zeta_N), \underline{h}(\underline{x})\underline{h}'(x)] < \delta$$

for some N. See Silvey (1980, pp. 35-36) for proof.

A well known application of the above procedure is Wynn's algorithm (see Wynn, 1970) to obtain D-optimal designs. In this situation ϕ satisfies

$$\phi[M(\zeta)] = \begin{cases} \log|M(\zeta)| & \text{if } M(\zeta) \text{ is nonsingular} \\ -\infty & \text{otherwise} \end{cases} .$$

Also $F_\phi[M(\zeta), \underline{h}(\underline{x})\underline{h}'(\underline{x})] = \underline{h}'(x)M^{-1}(\zeta)\underline{h}(\underline{x}) - m$ (see Silvey, 1980, p. 30).

In our application ϕ satisfies

$$\phi[M(\zeta)] = \begin{cases} \Lambda_2'[M(\zeta)] & \text{if } M_{XX}(\zeta) \text{ is nonsingular} \\ -\infty & \text{otherwise} \end{cases} .$$

Hence $\mathbf{M}' = \{M(\zeta): M_{XX}(\zeta) \text{ is nonsingular}\}$. An explicit expression for F_ϕ when $\phi = \Lambda_2'$ is derived in Theorem 3.3 and it is shown in Lemma 3.1 that Λ_2' satisfies the properties required to employ the sequential procedure described in subsection 3.5.1. The sequential procedure to obtain a Λ_2'-optimal design is given in subsection 3.5.2.

Theorem 3.3

Let $\zeta \in \mathbf{H}$. Then

$$F_{\Lambda_2}[M(\zeta), \underline{h}(x)\underline{h}'(\underline{x})] = \mathrm{tr}[T^{-1}L\{I_r \otimes ((\underline{b}(x) - \underline{v}(\underline{x},\zeta))(\underline{b}'(\underline{x}) - \underline{v}'(\underline{x},\zeta)))\}L'] - \Lambda_2'[M(\zeta)]$$

where $\underline{v}(\underline{x}, \zeta) = M_{ZX}(\zeta)M_{XX}^{-1}(\zeta)\underline{a}(x)$.

Recall that $\underline{a}'(\underline{x}_u)$ and $\underline{b}'(\underline{x}_u)$ are the row vectors which represent the u^{th} rows of X_0 and Z_0 respectively, and $\underline{h}'(\underline{x}) = [\underline{a}'(\underline{x}) : \underline{b}'(\underline{x})]$. The proof is given in Appendix G.

Lemma 3.1

Λ_2' is concave on **M**, differentiable on **M'** and a Λ_2'-optimal measure exists. The proof is given in Appendix H.

3.5.2 The Sequential Procedure to Obtain a Λ_2'-Optimal Design

1. Start with an initial design D_{N_0} (consisting of N_0 points) for which $X_0'X_0$ is nonsingular.

2. Obtain the design point $\underline{x}_{N_0+1}$ at which
$$\max_{\underline{x} \in X} F_{\Lambda_2'} [M(\zeta_{N_0}), \underline{h}(\underline{x})\underline{h}'(\underline{x})] \text{ is attained.}$$
The formulas necessary to calculate $F_{\Lambda_2'} [M(\zeta_{N_0}), \underline{h}(\underline{x})\underline{h}'(\underline{x})]$ and $M(\zeta_{N_0})$ are given in Theorem 3.3 and (3.24) respectively.

3. Obtain D_{N_0+1} (hence ζ_{N_0+1}) by augmenting D_{N_0} with $\underline{x}_{N_0+1}$. Recall that ζ_{N_0+1} is the design measure obtained by assigning probability $1/(N_0+1)$ to each design point in D_{N_0+1}.

4. Repeat this process to find $\underline{x}_{N_0+2}$, $\underline{x}_{N_0+3}$, ..., by replacing ζ_{N_0}, by ζ_{N_0+1}, ζ_{N_0+2}, ... until
$$\max_{\underline{x} \in X} F_{\Lambda_2'} [M(\zeta_N), \underline{h}(\underline{x})\underline{h}'(\underline{x})] < \delta$$

for some N and δ, where δ is a small positive number chosen a priori.

3.6 A Numerical Example

We now give an example to illustrate the sequential procedure described in subsection 3.5.1. In this example we have three responses and two controllable variables, x_1 and x_2, both in the range of [-1, 1]. The fitted models are given below.

$$
\begin{aligned}
y_1 &= \beta_{10} + \beta_{11}x_1 + \beta_{12}x_2 + \varepsilon_1 \\
y_2 &= \beta_{20} + \beta_{21}x_1 + \beta_{22}x_2 + \varepsilon_2 \qquad (3.29) \\
y_3 &= \beta_{30} + \beta_{31}x_1 + \beta_{32}x_2 + \beta_{33}x_1x_2 + \varepsilon_3 \quad .
\end{aligned}
$$

The true models are all second degree complete with pure quadratic and cross product terms in x_1 and x_2. The matrix X is partitioned as $X = [X_1 \colon X_2 \colon X_3]$ with both X_1, X_2 having the column of ones and the columns corresponding to x_1 and x_2 while X_3 has the column of ones and the columns corresponding to x_1, x_2, and x_1x_2. Similarly the matrix Z is partitioned as $Z = [Z_1 \colon Z_2 \colon Z_3]$ with both Z_1, Z_2 having columns corresponding to x_1x_2, x_1^2 and x_2^2 while Z_3 has the columns corresponding to x_1^2 and x_2^2 only. The matrix X_0 whose columns form a basis for the

columns of X contains the column of ones and the columns corresponding to x_1, x_2 and x_1x_2. The matrix Z_0 contains the columns corresponding to x_1x_2, x_1^2 and x_2^2. At each iteration $F_{\Lambda_2'}$ is evaluated with the use of the expression derived in Theorem 3.3. The necessary quantities required for the evaluation of $F_{\Lambda_2'}$ are given below. The matrix T is calculated using (3.16). Since $q = 8$, $r = 3$ and $\rho_1 = 3$, $L = [I_8 \colon \underline{0}_{8\times 1}]C$ (see (3.15)), where C (see (3.14)) is calculated below. The matrices $M_{ZX}(\zeta_N)$ and $M_{XX}(\zeta_N)$ at the N^{th} iteration were calculated using (3.24). The vectors $\underline{a}(\underline{x})$ and $\underline{b}(\underline{x})$ in this example are $\underline{a}'(\underline{x}) = (1, x_1, x_2, x_1x_2)$ and $\underline{b}'(\underline{x}) = (x_1x_2, x_1^2, x_2^2)$. The matrices H_i, $i = 1, 2, 3$ (see (3.13)) are

$$H_1 = H_2 = I_3 \quad \text{and} \quad H_3 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix},$$

and the matrix C is of order 9 x 9, and is equal to (see (3.14))

$$\begin{bmatrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0
\end{bmatrix}$$

The initial design consisted of the design points (1, 1), (1, -1), (-1, 1), (-1, -1). For this design $X_0'X_0$ is nonsingular. The design points obtained sequentially are given in Table 3.1 and the value of δ chosen is .005.

Table 3.1
The Augmented Points of the Λ_2'-Optimal Design

N	$\max_{\underline{x}\in X} F_{\Lambda_2'}\{M(\zeta_{N-1}),\underline{h}(\underline{x})\underline{h}'(\underline{x})\}$	$\underline{x}_N$	$\Lambda_2'[M(\zeta_{N-1})]$
5	16.875	(0, 0)	0
6	8.100	(0, 0)	2.70
7	3.750	(0, 0)	3.75
8	1.378	(0, 0)	4.13
9	.003	(0, 0)	4.22

As seen from the above Table, $\max_{\underline{x}\varepsilon\chi} F_{\Lambda_2'} [M(\zeta_8), \underline{h}(\underline{x})\underline{h}'(\underline{x})]$ is approximately zero. Thus we conclude that an approximate Λ_2'-optimal (hence a Λ_2-optimal) design for this example is (1, 1), (1, -1), (-1, 1), (-1, -1) and five replicated observations at the origin.

As described in section 3.4 the main objective of constructing a Λ_2-optimal design is to increase the power of the multivariate lack of fit test. This is achieved in view of the following facts.

1. The power is an increasing function of the eigenvalues of the noncentrality parameters matrix Ω.
2. Λ_2-optimal designs are intended to increase the trace of Ω, and hence at least one of its eigenvalues.

CHAPTER FOUR
DESIGNS TO REDUCE THE EFFECT OF NONNORMALITY ON SOME TESTS ASSOCIATED WITH REGRESSION PARAMETERS

4.1 Introduction

The effect of nonnormality on multiresponse regression tests was studied by Mardia (1971). Following Box and Watson (1962), he showed that the null distribution of an appropriate test statistic could be approximated by a beta distribution with modified degrees of freedom. The only assumption made was that the distribution of the error matrix $\varepsilon' = [\underline{\varepsilon}_1 : \underline{\varepsilon}_2 : \ldots : \underline{\varepsilon}_N]_{r \times N}$ is symmetric in $\underline{\varepsilon}_1, \underline{\varepsilon}_2, \ldots, \underline{\varepsilon}_N$, where r is the number of responses and N is the number of design settings. This is true if the error vectors are i.i.d. Mardia also showed that the modified degrees of freedom could be made exact if a certain multivariate measure of kurtosis for the regression (controllable) variables (denoted by C_X) is zero. Since C_X depends on the design points chosen, it could be made zero by a suitable choice of design.

In this chapter, we investigate methods of design construction to minimize $|C_X|$. This is done by adapting the design constructing techniques developed by Khuri and Myers (1981).

4.2 Robustness Criterion

Suppose $\underline{y}_u$, u = 1, 2, ..., N is an rx1 vector which consists of the observations made on the r responses at the u^{th} design setting.

Consider the model

$$Y = XB + \varepsilon \tag{4.1}$$

where $Y = [\underline{y}_1 : \underline{y}_2 : \ldots : \underline{y}_N]'$, $X = [X_1 : X_2 : \ldots : X_r]$. Each X_i is an Nxp_i full column-rank matrix: its first column is equal to $\underline{1}$ (the vector of ones) and the remaining elements are known functions of the settings of the k controllable variables, such that the u^{th} row corresponds to the u^{th} design setting. Also $B = \text{diag}(\underline{\beta}_1, \underline{\beta}_2, \ldots, \underline{\beta}_r)$ where $\underline{\beta}_i$, i = 1, 2, ..., r is the vector of regression coefficients associated with the i^{th} response and $\varepsilon = [\underline{\varepsilon}_1, \underline{\varepsilon}_2, \ldots, \underline{\varepsilon}_N]'$ is a Nxr matrix of errors such that the vector $\underline{\varepsilon}_u$, u = 1, 2, ..., N has the mean vector 0 and the variance-covariance matrix Σ, where Σ is unknown.

Let the columns of $[\underline{1}: D]$ form a basis for the columns of X. Then D has order $Nx\rho$ and rank ρ, where $\text{rank}(X) = \rho + 1$. Also

$$X_i = [\underline{1}: D]E_i \quad , \quad i = 1, 2, \ldots, r \quad , \tag{4.2}$$

where E_i is a unique matrix of rank p_i. In view of (4.2) we can write (4.1) as

$$Y = [1 : D][E_1 : E_2 : \ldots : E_r]\begin{bmatrix} \underline{\beta}_1 & 0 & & 0 \\ & \underline{\beta}_2 & & 0 \\ & & \ddots & \\ 0 & 0 & & \beta_r \end{bmatrix} + \varepsilon \quad .$$

i.e.,

$$Y = [\underline{1} : D][E_1\underline{\beta}_1 : E_2\underline{\beta}_2 : \ldots : E_r\underline{\beta}_r] + \varepsilon \quad .$$

Equivalently

$$Y = \underline{1}\ \gamma_0' + DG + \varepsilon \tag{4.3}$$

where $\underline{\gamma}_0$ is an rx1 vector of parameters, $G = [\underline{\gamma}_1 : \underline{\gamma}_2 : \ldots : \underline{\gamma}_r]$ where γ_i, $i = 1, 2, \ldots, r$ is a px1 vector of parameters with some of its elements possibly equal to zero. Let A be a known rxm full column-rank matrix $(r > m)$ such that $GA = F$, where F is a pxm matrix of linear parametric functions. Then we can write (4.3) as

$$Y^a = 1(\underline{\gamma}_0^a)' + DF + \varepsilon^a \tag{4.4}$$

where $Y^a = YA$, $(\underline{\gamma}_0^a)' = \underline{\gamma}_0'A$ and $\varepsilon^a = \varepsilon A$. It can be easily seen that each column of $(\varepsilon^a)'$ has a zero mean vector and a variance-covariance matrix $A'\Sigma A$. Observe that the model (4.4) is of the same form as model (2.1) in Mardia (1971). Moreover, let $M = D(D'D)^{-1}D'$ and $Z = Y^a - \underline{1}(\underline{\bar{Y}}^a)'$

where $\bar{\underline{Y}}^a$ is the mx1 vector whose elements are the column means of Y^a. Then an appropriate test statistic to test the null hypothesis $H_0 : F = 0$ against $H_a: F \neq 0$ is , $V = tr(H(E + H)^{-1})$ where $H = Z'MZ$, $E = Z'(I - M)Z$ and 'tr' donotes trace. Let $s = \min(r, \rho)$, $P = r\rho/2$ and $Q = [s(N - 1) - r\rho]/2$. Then Mardia showed that, even under nonnormal situations the null distribution of V/s is an approximate beta distribution with the parameters δP and δQ, provided the joint probability density function of $(\varepsilon^a)' = [\underline{\varepsilon}_1^a : \underline{\varepsilon}_2^a \cdots : \underline{\varepsilon}_N^a]$ is symmetric in $\underline{\varepsilon}_1^a, \underline{\varepsilon}_2^a, \ldots, \underline{\varepsilon}_N^a$. The quantity δ is called the corrective factor and is equal to one if $C_X = 0$ where

$$C_X = \{N(N - 1)(N + 1)/[\rho(N - \rho - 1)(N - 3)]\} \times \{d - [\rho(\rho + 2)(N - 1)/N(N + 1)]\} \quad . \tag{4.5}$$

In (4.5), $d = \sum_{i=1}^{N} d_{ii}^2$ and d_{ii} is the i^{th} diagonal element of M. Now if

$$g(D) = d - \rho(\rho + 2)(N - 1)/[N(N + 1)] \quad , \tag{4.6}$$

then $g(D) = 0$ implies $C_X = 0$. In this case $\delta = 1$ and hence the approximate distribution of V/s is identical to the distribution obtained under the normality assumption. Now since d depends on the design points, an appropriate design will minimize $|g(D)|$. Such a design is said to satisfy the robustness criterion.

The quantities C_X and g(D) above are identical to the corresponding quantities in the single response model (see Khuri and Myers, 1981). Therefore, multiresponse designs which satisfy the robustness criterion can be constructed by employing the same technique developed for the single response case by Khuri and Myers (1981).

4.3 The Construction of a Design Satisfying the Robustness Criterion for a Given Number of Experimental Runs

In this section we describe the Khuri and Myers technique for the single response case, and as mentioned earlier, the same technique works for the multiresponse case.

Box and Watson (1962) have shown that d in (4.5) is invariant under any nonsingular transformation of the form S = DT of the columns of D. Therefore, the columns of D can be regarded as orthogonal. Suppose $D = [D_1' : 0'_{n_0 x\rho}]'$, where D_1 is the submatrix of D consisting of the $n = N - n_0$ noncenter points and the controllable variables are scaled so that $D_1'D_1 = nI$. Let $C = [c_{ij}]$ be an orthogonal matrix of order nxn whose first row consists of elements equal to $(1/n)^{1/2}$, and the next ρ rows ($\rho < n$) are the corresponding columns of $D_1/(n)^{1/2}$. Then

$$g(D) = \sum_{j=1}^{n} \left(\sum_{i=2}^{\rho+1} c_{ij}^2\right)^2 - \rho(\rho + 2)(N - 1)/[N(N + 1)] \quad . \qquad (4.7)$$

(See Khuri and Myers (1981) for details.) In addition, the elements of C satisfy

$$\sum_{j=1}^{n} c_{ij}^2 = 1 \quad , \quad \sum_{j=1}^{n} c_{ij} = 0 \quad , \quad i = 2, 3, \ldots, \rho + 1 \quad ,$$

and (4.8)

$$\sum_{j=1}^{n} c_{ij}c_{lj} = 0 \quad , \quad 2 \le i \le l \le \rho + 1 \quad .$$

Khuri and Myers used the following approach to obtain a design which minimizes $|g(D)|$, or equivalently $g^2(D)$, subject to the constraints given by (4.8). Since $g^2(D)$ is continuous, and the domain of $\{c_{ij}: \ 2 \le i \le \rho + 1, 1 \le j \le n\}$ satisfying (4.8) is a closed and bounded subset in the Euclidean space, the absolute minimum of $g^2(D)$ must exist, and be attained in that domain. If a solution to $g(D) = 0$ exists, then the absolute minimum of $g^2(D)$ must be zero, and vice versa. The method of Lagrange multipliers can, therefore, be used as follows to obtain a solution for $\{c_{ij}: 2 \le i \le \rho + 1, 1 \le j \le n\}$, which leads to a design satisfying the robustness criterion. First consider the Lagrange function

$$L = g^2(D) + \sum_{i=2}^{\rho+1} \lambda_i \left(\sum_{j=1}^{n} c_{ij}^2 - 1\right) + \sum_{i=2}^{\rho+1} \mu_i \sum_{j=1}^{n} c_{ij}$$

(4.9)

$$+ \sum_{i<l=3}^{\rho+1} \eta_{il} \sum_{j=1}^{n} c_{ij}c_{lj}$$

where λ_i, μ_i, $i = 2, 3, \ldots, \rho + 1$, and η_{il}, $2 \le i < l \le \rho + 1$ are Lagrange multipliers. Then solve the $(2\rho + \rho(\rho - 1)/2)$ equations in

(4.8), and the $n\rho$ equations $\frac{\partial L}{\partial c_{ij}} = 0$, $i = 2, 3, \ldots, \rho + 1$; $j = 1, 2, \ldots, n$ for c_{ij} and the Lagrange multipliers.

4.4 Finding an Optimal Number of Center Points to Construct an Orthogonal Central Composite Design Satisfying the Robustness Criterion

A central composite design is obtained by adding the following design points to a 2^k factorial design, or a fractional factorial design (see Myers, 1971, p. 127).

x_1	x_2	x_3	$\cdots$	x_k
0	0	0	$\cdots$	0
$\vdots$	$\vdots$	$\vdots$	$\cdots$	0
0	0	0	$\cdots$	0
$-\alpha$	0	0	$\cdots$	0
α	0	0	$\cdots$	0
0	$-\alpha$	0	$\cdots$	0
0	α	0	$\cdots$	0
$\vdots$	$\vdots$	$\vdots$	$\cdots$	$\vdots$
0	0	0	$\cdots$	$-\alpha$
0	0	0	$\cdots$	α

Thus if there are n_0 center points, then a central composite design which contains a complete factorial design, has $2^k + n_0 + 2k$ design points.

Suppose the matrix D in (4.2) corresponds to the design matrix of an orthogonal central composite design, which contains a complete factorial design. Then

$$g(D) = \bar{\Lambda}[(k/(\bar{\Lambda} + 2\alpha^2)) + (k(1 - C')^2/\phi + (k(k - 1)/2\bar{\Lambda})]^2 + n_0 k^2 (C')^4/\phi$$

$$+ 2k[(\alpha^2/(\bar{\Lambda} + 2\alpha^2)) + ((\alpha^2 - C')^2/\phi) + ((k - 1)(C')^2/\phi)]^2$$

$$-[\rho(\rho + 2)(N - 1)/N(N + 1)] \quad .$$

Here, $\bar{\Lambda} = 2^k$, $\phi = (\bar{\Lambda}T - 4\bar{\Lambda}\alpha^4)/(\bar{\Lambda} + T)$, $C' = (\bar{\Lambda} + 2\alpha^2)/(\bar{\Lambda} + T)$, $\rho = (3k + k^2)/2$, and $\alpha = (\nu\bar{\Lambda}/4)^{1/4}$ where $T = 2k + n_0$, $N = T + \bar{\Lambda}$, and $\nu = ((\bar{\Lambda} + T)^{1/2} - \bar{\Lambda}^{1/2})^2$. Note that this value of α guarantees the orthogonality of the design (see Myers, 1971, Ch. 7). Now g(D) is a function of n_0 (the number of center points). Thus an optimal value of n_0 can be obtained by simply plotting g(D) for different values of n_0. However, it is possible that the g(D) value corresponding to the optimal value of n_0 may not be close to zero.

4.5 A Numerical Example

This is a very simple example to illustrate the procedure in Section 4.3. We have three responses dependent on two controllable variables x_1 and x_2. The regression models are

$$
\begin{aligned}
y_1 &= \beta_{10} + \beta_{11}x_1 + \beta_{12}x_2 + \varepsilon_1 \\
y_2 &= \beta_{20} + \beta_{21}\,x_1 + \varepsilon_2 \\
y_3 &= \beta_{30} + \beta_{31}\,x_1 + \beta_{32}x_2 + \varepsilon_3 \quad .
\end{aligned}
\tag{4.10}
$$

Then each of X_1 and X_3 contains the column of ones and the columns corresponding to x_1, x_2, X_2 contains the column of ones and the column corresponding to x_1. Therefore D in (4.2) contains the columns corresponding to x_1, x_2 and $\rho = 2$. Also in (4.2)

$$E_1 = E_3 = I_3 \quad \text{and} \quad E_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \quad .$$

In (4.3),

$$G = \begin{bmatrix} \beta_{11} & \beta_{21} & \beta_{31} \\ \beta_{12} & 0 & \beta_{32} \end{bmatrix} \quad .$$

We are interested in testing the null hypothesis

$$H_0: \; F = \underset{\sim}{0}_{2\times 1} \quad \text{against} \quad H_a: \; F \neq \underset{\sim}{0}_{2\times 1}$$

where $F = (\beta_{11} - \beta_{31}, \beta_{12} - \beta_{32})'$. Then the matrix A which satisifes $F = GA$ is $A = (1, 0, -1)'$. For $n = 4$, the values of c_{ij} $i = 2, 3, \ldots, p+1$; $j = 1, 2, \ldots, n$ and the Lagrange multipliers satisfying $\frac{\partial L}{\partial c_{ij}} = 0$ and (4.8) were obtained by using the subroutine ZSPOW in FORTRAN. This subroutine solves a system of nonlinear equations. The second and third rows of C are given by rows one and two of the matrix

$$C_1 = \begin{bmatrix} .500 & .500 & -.499 & -.500 \\ .499 & -.499 & .500 & -.500 \end{bmatrix} .$$

Therefore

$$D' = (4)^{1/2} C_1 = \begin{bmatrix} 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \end{bmatrix} \quad \text{or} \quad D = \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ -1 & 1 \\ -1 & -1 \end{bmatrix}$$

This is a 2^2 factorial design.

CHAPTER FIVE
CONCLUSIONS

This dissertation has been primarily concerned with the construction of designs for linear multiresponse models. We considered three different topics in the area, which are important for better design of multiresponse experiments.

In Chapter Two, D-optimal designs for multiresponse experiments were discussed, and a sequential procedure was developed to obtain multiresponse D-optimal designs when Σ, the variance-covariance matrix of the error vector is not known. However, this procedure cannot be used to generate multiresponse D-optimal designs for a fixed number of design points. It would be desirable to find a procedure that could be used to generate multiresponse D-optimal designs in such situations.

In Chapter Three, two design criteria were developed to increase the power of the multiresponse lack of fit test. They are the multivariate extensions of the Λ_1 and Λ_2 criteria proposed by Jones and Mitchell (1978) for the single response case. A sequential procedure to obtain Λ_2-optimal designs was also presented. The Λ_1 and Λ_2 criteria increase the trace of Ω [the noncentrality parameter matrix associated with the distribution of G_1 (see (3.8))], and hence the power of the multiresponse lack of fit test which is an increasing function of the

eigenvalues of Ω. The question of whether it is possible to develop criteria to increase the smallest eigenvalue of Ω should be investigated in the future.

In Chapter Four we considered the topic of constructing designs to reduce the effect of nonnormality of the error distribution on tests of hypotheses associated with linear combinations of the parameter vectors. A related topic for future research is the construction of designs to reduce the effect of nonnormality of the error distribution on the multiresponse test of lack of fit.

APPENDIX A
DEFINITIONS RELATED TO CONVERGENCE OF RANDOM VECTORS AND RANDOM MATRICES

Let (Ω, Σ, P) be a probability space, and $\underline{x}$, $\{\underline{x}_N, N \geq 1\}$ be random vectors defined on Ω into R^k.

Definition A.1

$\underline{x}_N$ is said to converge in probability to $\underline{x}$, or equivalently, $\underline{x}_N$ is a consistent estimator of $\underline{x}$ if for any $\varepsilon > 0$

$$\lim_{N\to\infty} P\{\omega: \omega\varepsilon\Omega \text{ and } ||\underline{x}_N(\omega) - \underline{x}(\omega)|| > \varepsilon\} = 0$$

where $|| \ ||$ denotes the Euclidean norm in R^k.

Definition A.2

$\underline{x}_N$ is said to converge with probability 1 (ω.p.1) to $\underline{x}$ if

$$P\{\omega: \omega\varepsilon\Omega \text{ and } \lim_{N\to\infty} \underline{x}_N(\omega) = \underline{x}(\omega)\} = 1.$$

A matrix S of finite order is said to be a random matrix if each of its elements is a random variable on Ω.

Definition A.3

Let S and $\{S_N, N > 1\}$ be matrices of the same order. Then S_N is said to be a consistent estimator of S if and only if s_{ij}^N is a consistent estimator of s_{ij} for all i, j where $S_N = [s_{ij}^N]$ and $S = [s_{ij}]$.

Suppose S and $\{S_N, N > 1\}$ are symmetric matrices, $\underline{b}(S)$ is the vector which consists of elements above or on the diagonal of S, and if e_N is the Euclidean distance between $\underline{b}(S_N)$ and $\underline{b}(S)$, then the consistency of S_N is equivalent to convergence of e_N to 0 in probability.

APPENDIX B
RESULTS IN MATRIX ALGEBRA

Theorem B.1

Let T be a non-singular matrix of order nxn. Then

1. The determinant of T denoted $|T|$ is a polynomial in its elements.

2. Each element of T^{-1} is a rational function of elements of T.

Proof.

(1) Using the definition given in Aitken (1951, p. 30) we have

$$|T| = \Sigma \pm t_{1\alpha} t_{2\beta} \dots t_{n\nu} \tag{B.1}$$

with the summation of n! terms being extended over all permutations $(\alpha, \beta, \dots, \nu)$ of column subscripts of the elements t_{ij} of T and the sign + or - being prefixed to any term according as the permutation is even or odd, respectively. It is clear that every term in the summation of (B.1) is a polynomial in the set of elements $\{t_{ij}, 1 < i, j < n\}$, and since a finite sum of polynomials is itself a polynomial, we have the required result.

(2) We have from Lancaster (1969, p. 36)

$$T^{-1} = (\text{Adj } T)/|T|$$

where $(\text{Adj}T)_{ij} = (-1)^{i+j} M_{ji}$ and M_{ji} is the determinant of the matrix obtained by deleting the j^{th} row and the i^{th} column. Applying (1) for this matrix we have that $(\text{Adj}T)_{ij}$ is a polynomial in elements of T and since $|T|$ is a polynomial in elements of T we obtain the required result.

Theorem B.2

Let T^{ij}, $1 \le i < j \le n$ be a symmetric matrix of order nxn such that its ij^{th} and ji^{th} elements are equal to 1 and all other elements are equal to zero. Then the non-zero eigenvalues of T^{ij} are 1 and -1 for all $1 \le i < j \le n$.

Proof.

$$T^{ij} = \begin{array}{c@{}c} & \begin{array}{ccccccccc} & & & i & & j & & \end{array} \\ \begin{array}{c} \\ \\ i \\ \\ j \\ \\ \end{array} & \left[\begin{array}{cccccccc} 0 & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots & & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots & & \vdots & & \vdots \\ 0 & \cdots & & 1 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots & & \vdots & & \vdots \\ 0 & \cdots & & 0 & \cdots & 0 & \cdots & 0 \end{array}\right] \end{array} .$$

Since n-2 columns of T are identically equal to zero, $\mathrm{rank}(T^{ij}) = 2$. Therefore T has only two non zero eigenvalues. Let $\underline{p}^{ij}$ be a nx1 vector such that its i^{th} and j^{th} elements are equal to 1 and all the other elements are equal to zero. Then $\underline{p}^{ij}$ is an eigenvector associated with 1. For let

$$\underline{q}^{ij} = T^{ij}\, p^{ij} \quad .$$

Then

$$q_u^{ij} = \sum_{k=1}^{n} t_{uk}^{ij}\, p_k^{ij},\ 1 \leq u \leq n \text{ where } \underline{q}^{ij} = (q_1^{ij} \ \ldots \ q_n^{ij})' \quad ,$$

$$T^{ij} = [t_{uv}^{ij}], \text{ and } \underline{p}^{ij} = (p_1^{ij}, p_2^{ij}, \ldots, p_n^{ij})' \quad .$$

Therefore

$$q_i^{ij} = t_{ij}^{ij}\, p_j^{ij} = 1$$

$$q_j^{ij} = t_{ji}^{ij}\, p_i^{ij} = 1$$

$$q_u^{ij} = 0 \qquad \text{for } u \neq i, j \quad \text{and } 1 \leq u \leq n \quad .$$

Thus $\underline{q}^{ij} = p^{ij}$.

Therefore $T^{ij}\underline{p}^{ij} = \underline{p}^{ij}$.

Also it is clear that $tr(T^{ij}) = 0$, since all the diagonal elements of T are zero. Thus if λ is the other nonzero eigenvalue then $1 + \lambda = 0$, which implies that $\lambda = -1$.

Theorem B.3

Let S be a pxr matrix and T be an rxr symmetric matrix. Then

$$e_{min}(T)\, tr(SS') < tr(STS') < e_{max}(T) tr(SS')$$

where e_{min} and e_{max} denote the smallest and the largest eigenvalues of a matrix.

Proof.

$$tr(STS') = tr[S(T - e_{max}(T)I_r + e_{max}(T)I_r)S']$$

$$= tr[-S(e_{max}(T)I_r - T)S' + e_{max}(T)SS']$$

$$< e_{max}(T)tr(SS')$$

(since $e_{max}(T)I_r - T$ is positive semidefinite).

Similarly

$$tr(STS') = tr[S(T - e_{min}(T)I_r + e_{min}(T)I_r)S']$$

$$= tr[S(T - e_{min}(T)I_r)S'] + e_{min}(T)tr(SS')$$

$$> e_{min}(T)tr(SS') \quad .$$

Therefore

$$e_{min}(T)tr(SS') < tr(STS') < e_{max}(T)tr(SS') \quad .$$

Corollary B.1

Let A and B be two square matrices of the same order. If A is symmetric and B is positive semidefinite, then

$$tr(AB) < e_{max}(A)tr(B) \quad .$$

Proof.

Since B is positive semidefinite $B = S'S$ for some matrix S. Then $tr(AB) = tr(AS'S) = tr(SAS') < e_{max}(A)tr(SS') = e_{max}(A)tr(B)$ using the above theorem.

APPENDIX C
CONSISTENCY OF $\hat{\Sigma}_N$

By definition

$$N\hat{\sigma}^N_{ij} = (\underline{Y}_i - F_i\underline{\tilde{\beta}}_i)'(\underline{Y}_j - F_j\underline{\tilde{\beta}}_j)$$

$$= (\underline{Y}_i - R_i\underline{Y}_i)'(\underline{Y}_j - R_j\underline{Y}_j) \quad \text{where } R_i = F_i(F_i'F_i)^{-1}F_i'$$

$$i = 1, 2, \ldots, r$$

$$= \underline{Y}_i'(I_N - R_i)'(I_N - R_j)\underline{Y}_j$$

$$= (\underline{\beta}_i'F_i' + \underline{\varepsilon}_i')(I_N - R_i)'(I_N - R_j)(F_j\underline{\beta}_j + \underline{\varepsilon}_j)$$

$$= \underline{\varepsilon}_i'(I_N - R_i)(I_N - R_j)\underline{\varepsilon}_j \tag{C.1}$$

$$= \sum_{u=1}^{N}\sum_{v=1}^{N} \varepsilon_{ui}t^{ij}_{uv}\varepsilon_{vj} \quad \text{where } T^{ij} = (I_N - R_i)(I_N - R_j)$$

$$\text{and } T^{ij} = [t^{ij}_{uv}] \quad .$$

Therefore

$$E(N\hat{\sigma}_{ij}^{N}) = \sum_{u=1}^{N} t_{uu}^{ij}E(\varepsilon_{ui}\varepsilon_{uj}) + \sum_{\substack{u=1 \\ u\neq v}}^{N} \sum_{v=1}^{N} t_{uv}^{ij}E(\varepsilon_{ui}\varepsilon_{vj})$$

$$= \sum_{u=1}^{N} t_{uu}^{ij}\sigma_{ij} + 0 \quad (\text{since } E(\varepsilon_{ui}\varepsilon_{vj}) = 0, \text{ for } u\neq v)$$

$$= \sigma_{ij}tr(T^{ij}) \quad .$$

Now

$$tr(T^{ij}) = tr[I_N - R_i - R_j + R_iR_j]$$

$$= N - p_i - p_j + tr(R_iR_j) \quad .$$

Therefore

$$E[N\hat{\sigma}_{ij}^{N}/(N - p_i - p_j + tr(R_iR_j)] = \sigma_{ij} \quad .$$

It is clear from (C.1) that $\hat{\sigma}_{ij}^{N}$ is a bilinear form. Therefore using (52) in Searle (1971, p. 65) which is:

$$\mathrm{Var}(\underline{x}_1' A_{12} \underline{x}_2) = \mathrm{tr}(A_{12}C_{21})^2 + \mathrm{tr}(A_{12}C_{22}A_{21}C_{11}) + \underline{\mu}_1' A_{12}C_{22}A_{21}\underline{\mu}_1$$

$$+ \underline{\mu}_2' A_{21}C_{11}A_{12}\underline{\mu}_2 + 2\underline{\mu}_1' A_{12}C_{21}A_{12}\underline{\mu}_2 \quad ,$$

with

$$\underline{x}_1' = \underline{\varepsilon}_i' \quad , \quad \underline{x}_2 = \underline{\varepsilon}_j \quad , \quad A_{12} = (I_N - R_i)(I_N - R_j) \quad ,$$

$$C_{11} = \sigma_{ii} I_N, \quad C_{22} = \sigma_{jj} I_N \quad , \quad C_{12} = \sigma_{ij} I_N \quad , \quad \underline{\mu}_1 = \underline{\mu}_2 = 0$$

we have

$$\mathrm{Var}[\underline{\varepsilon}_i'(I_N - R_i)(I_N - R_j)\underline{\varepsilon}_j] = \sigma_{ij}^2 \; \mathrm{tr}[(I_N - R_i)(I_N - R_j)]^2$$

$$+ \sigma_{ii}\sigma_{jj}\mathrm{tr}[(I_N - R_i)(I_N - R_j)(I_N - R_j)(I_N - R_i)]$$

$$= \sigma_{ij}^2 \mathrm{tr}[(I_N - R_i)(I_N - R_j)]^2 + \sigma_{ii}\sigma_{jj}\mathrm{tr}[(I_N - R_i)(I_N - R_j)] \quad .$$

We also use Corollary B.1 in Appendix B to obtain

$$\mathrm{tr}[(I_N - R_i)(I_N - R_j)] < e_{max}(I_N - R_i)\mathrm{tr}(I_N - R_j) = N - p_j \quad ,$$

(since $I_N - R_i$ is idempotent, $e_{max}(I_N - R_i) = 1$) .

Moreover

$$tr[(I_N - R_i)(I_N - R_j)]^2 = tr[I_N - R_i - R_j + R_i R_j R_i R_j]$$

$$= N - p_i - p_j + tr(R_i R_j R_i R_j)$$

$$\leq N - p_i - p_j + e_{max}(R_i) tr(R_j R_i R_j)$$

$$= N - p_i - p_j + tr(R_j R_i R_j) \quad .$$

Also, since $(I_N - R_i)$ is positive semidefinite,

$R_j R_j - R_j R_i R_j$ is postive semidefinite.

Therefore

$$tr(R_j R_i R_j) \leq tr(R_j R_j)$$

$$= tr(R_j)$$

$$= p_j$$

$$< p_j^2 \quad , \quad \text{since } p_j > 1 \quad .$$

Therefore

$$tr[(I_N - R_i)(I_N - R_j)]^2 < N - p_i - p_j + p_j^2 \quad ,$$

and

$$Var[\underline{\varepsilon}_i'(I_N - R_i)(I_N - R_j)\underline{\varepsilon}_j] < \sigma_{ij}^2(N - p_i - p_j + p_j^2) + \sigma_{ii}\sigma_{jj}(N - p_j) \quad .$$

Hence

$$Var(N\hat{\sigma}_{ij}^N) < \sigma_{ij}^2(N - p_i - p_j + p_j^2) + \sigma_{ii}\sigma_{jj}(N - p_j) \quad .$$

Let $\omega\varepsilon\Omega$ and $\varepsilon > 0$. Then from the Chebyshev inequality

$$P\{\omega: |N\hat{\sigma}_{ij}^N(\omega)/((N - p_i - p_j + trR_iR_j) - \sigma_{ij})| > \varepsilon\}$$

$$< [\sigma_{ij}^2(N - p_i - p_j + p_j^2) + \sigma_{ii}\sigma_{jj}(N - p_j)]/(N - p_i - p_j + tr(R_iR_j))^2\varepsilon^2 \ .$$

The term on the r.h.s. $\to 0$ as $N \to \infty$. Therefore $N\hat{\sigma}_{ij}^{N}/(N - p_i + p_j + \mathrm{tr}(R_iR_j))$ converges in probability to σ_{ij}, $1 \leq i, j \leq r$. Also $(N - p_i - p_j + \mathrm{tr}(R_iR_j))/N$ converges to 1, since by Corollary B.1 $\mathrm{tr}(R_iR_j) \leq \mathrm{tr}(R_j) = p_j$. Thus $\hat{\sigma}_{ij}^{N}$ converges in probability to σ_{ij}, $1 \leq i, j \leq r$. Equivalently $\hat{\Sigma}_N$ is a consistent estimator of Σ.

APPENDIX D
PROOF OF THEOREM 2.3

Theorem 2.3

$$|M(\zeta, \Sigma)| = \prod_{i=1}^{r} (\sigma^{ii})^{p_i} |M(\zeta, A^{-1})], \ \zeta \in H \text{ and } \Sigma^{-1} = [\sigma^{ij}] \quad .$$

Proof. From (2.14) we have

$$M(\zeta, \Sigma) = \sum_{j=1}^{s} \lambda_j \phi(\underline{x}_j) \Sigma^{-1} \phi'(\underline{x}_j) \quad \text{where} \quad 0 < \lambda_j < 1 \quad ,$$

$$1 < s < p', \ p' = p(p+1)/2 + 1 \quad , \quad \text{and} \ \sum_{j=1}^{s} \lambda_j = 1 \quad .$$

Define

$$X_i = [\lambda_1^{1/2} \underline{f}_i(\underline{x}_1), \lambda_2^{1/2} \underline{f}_i(\underline{x}_2), \ldots, \lambda_s^{1/2} \underline{f}_i(\underline{x}_s)]'_{\ p_i x s}, \ i = 1, 2, \ldots, r \quad .$$

and let

$$X = \begin{bmatrix} X_1 & 0 & & 0 \\ & X_2 & & \\ & & \ddots & \\ 0 & & & X_r \end{bmatrix}_{srxp} \quad .$$

Then $X'(\Sigma^{-1} \otimes I)]X =$

$$\begin{bmatrix} X_1' & 0 & \cdots & 0 \\ 0 & X_2' & \cdots & 0 \\ 0 & 0 & \cdots & X_r' \end{bmatrix} (\Sigma^{-1} \otimes I_s) \begin{bmatrix} X_1 & 0 & \cdots & 0 \\ 0 & X_2 & \cdots & 0 \\ 0 & 0 & \cdots & X_r \end{bmatrix}$$

$$= \begin{bmatrix} \sigma^{11} X_1' X_1 & \cdots & \sigma^{1r} X_1' X_r \\ & & \vdots \\ \sigma^{r1} X_r' X_1 & \cdots & \sigma^{rr} X_r' X_r \end{bmatrix}.$$

$$= \begin{bmatrix} \sigma^{11} \sum_{j=1}^{s} \lambda_j \underline{f}_1(\underline{x}_j) \underline{f}_1'(\underline{x}_j) & \cdots & \sigma^{1r} \sum_{j=1}^{s} \lambda_j \underline{f}_1(\underline{x}_j) \underline{f}_r'(\underline{x}_j) \\ \vdots & & \vdots \\ \sigma^{1r} \sum_{j=1}^{s} \lambda_j \underline{f}_r(\underline{x}_j) f_1'(\underline{x}_j) & & \sigma^{rr} \sum_{j=1}^{s} \lambda_j \underline{f}_r(\underline{x}_j) \underline{f}_r'(\underline{x}_j) \end{bmatrix}$$

$$= \sum_{j=1}^{s} \lambda_j \begin{bmatrix} \sigma^{11} \underline{f}_1(\underline{x}_j) \underline{f}_1'(\underline{x}_j) & \cdots & \sigma^{1r} \underline{f}_1(\underline{x}_j) f_r'(\underline{x}_j) \\ \vdots & & \vdots \\ \sigma^{1r} \underline{f}_r(\underline{x}_j) \underline{f}_1'(\underline{x}_j) & \cdots & \sigma^{rr} \underline{f}_r(\underline{x}_j) \underline{f}_r'(\underline{x}_j) \end{bmatrix}$$

$$= \sum_{j=1}^{s} \lambda_j \phi(\underline{x}_j)\Sigma^{-1}\phi'(\underline{x}_j) \quad , \quad \text{where} \quad \phi(\underline{x}_j) = \begin{bmatrix} \underline{f}_1(\underline{x}_j) & & \\ & \ddots & \\ & & \underline{f}_r(\underline{x}_j) \end{bmatrix}$$

$$= M(\zeta, \Sigma), \quad \text{by formula (2.14).}$$

Therefore

$$|M(\zeta, \Sigma)| = |X'(\Sigma^{-1} \otimes I_s)X| \quad .$$

Let $S = \begin{bmatrix} \sigma^{11} & & \\ & \ddots & \\ & & \sigma^{rr} \end{bmatrix}$.

Then

$$A = S^{-1/2}\,\Sigma^{-1}S^{-1/2}$$

and

$$\Sigma^{-1} \otimes I_s = S^{1/2}AS^{1/2} \otimes I_s$$

$$= (S^{1/2} \otimes I_s)(A \otimes I_s)(S^{1/2} \otimes I_s)$$

$$|X'(\Sigma^{-1} \otimes I_s)X| = |X'(S^{1/2} \otimes I_s)(A \otimes I_s)(S^{1/2} \otimes I_s)X|$$

$$= |CX'(A \otimes I_s)XC| \quad \text{where}$$

$$C = \begin{bmatrix} (\sigma^{11})^{1/2} I_{p_1} & & & \\ & (\sigma^{22})^{1/2} I_{p_2} & & \\ & & \ddots & \\ & & & (\sigma^{rr})^{1/2} I_{p_r} \end{bmatrix} .$$

Hence

$$|X'(\Sigma^{-1} \otimes I_s)X| = |C^2||X'(A \otimes I_s)X|$$

$$= \prod_{i=1}^{r} (\sigma^{ii})^{p_i} |X'(A \otimes I_s)X|$$

$$= \prod_{i=1}^{r} (\sigma^{ii})^{p_i} |M(\zeta, A^{-1})| \quad .$$

APPENDIX E
CONSISTENCY OF $\hat{A}_N$

Proof

By definition

$$A = (\text{diag } \Sigma^{-1})^{-1/2}\, \Sigma^{-1} (\text{diag } \Sigma^{-1})^{-1/2}$$

where $\Sigma = [\sigma_{ij}]$, $1 \leq i, j \leq r$.

Suppose $\Sigma^{-1} = [\sigma^{ij}]$ then

$$a_{ij} = \sigma^{ij}/(\sigma^{ii}\sigma^{jj})^{1/2} \qquad 1 \leq i, j \leq r \quad .$$

Since $\sigma^{ii} > 0$, $1 \leq i \leq r$, a_{ij} is well defined for all $1 \leq i, j \leq r$. Using Theorem B.1(2) we have that σ^{ij} is a rational function of σ_{11}, σ_{12}, ..., σ_{1r} and hence each a_{ij}. $1 \leq i, j \leq r$ is a well defined rational function of σ_{11}, σ_{12}, ..., σ_{rr}. Thus, each a_{ij} is a continuous function of σ_{11}, σ_{12}, ..., σ_{rr}. Since $\hat{\Sigma}_N$ is a consistent estimator

of Σ and consistency is preserved under continuous transformations, it follows that $\hat{a}_{ij}^{N}$ is a consistent estimator of a_{ij}, $1 \leq i < j \leq r$ where $\hat{A}_N = [\hat{a}_{ij}^{N}]$. Equivalently $\hat{A}_N$ is a consistent estimator of A.

APPENDIX F
PROOF OF THEOREM 2.4

We begin by introducing some notation and definitions related to Theorem 2.4 and then establish the lemmas that are required for its proof.

F.1 Notation

Let $\Lambda(\Sigma)$ be the set of all matrices of the form $M(\zeta, \Sigma)$ where $\zeta \varepsilon H$, the set of design measures defined on χ. Then from (2.14) we have

$$M(\zeta, \Sigma) = \sum_{u=1}^{s} \lambda_u \phi(\underline{x}_u) \Sigma^{-1} \phi'(\underline{x}_u) \tag{F.1}$$

where $0 < s < (p(p + 1)/2) + 1$, $0 < \lambda_u < 1$, and $\sum_{u=1}^{s} \lambda_u = 1$.

This can be rewritten as

$$M(\zeta, \Sigma) = \sum_{u=1}^{p'} \lambda_u \phi(\underline{x}_u) \Sigma^{-1} \phi'(\underline{x}_u) \tag{F.2}$$

where $p' = (p(p + 1)/2) + 1$ and $0 < \lambda_u < 1$ with $\sum_{u=1}^{p'} \lambda_u = 1$. (Observe that $\lambda_u = 0$ for $s < u \leqslant p'$).

Let **A** be as in Section 2.6 and let **A'** be the subset of **A** which consists of all positive definite matrices in **A**. Then $\Lambda(A)$ is defined to be the set of all matrices of the form $M(\zeta, A^{-1})$, $A \varepsilon \mathbf{A'}$ given by

$$M(\zeta, A^{-1}) = \sum_{u=1}^{p'} \lambda_u \phi(\underline{x}_u) A \phi'(\underline{x}_u) \tag{F.3}$$

where $0 < \lambda_u < 1$, and $\sum_{u=1}^{p'} \lambda_u = 1$.

Observe that (F.3) was obtained on replacing Σ^{-1} by A in (F.2), and that $\Lambda(A)$ is a convex set.

Let $\{\underline{x}_N, N > 1\}$ be a sequence of design points in χ and $\{\zeta_N, N > 1\}$ be a sequence of discrete designs such that the spectrum of ζ_{N+1} is the union of the spectrum of ζ_N and $\underline{x}_{N+1}$. If $N' > 1$ and $\zeta_{N'}$ is a design such that $M(\zeta_{N'}, A^{-1}) \varepsilon \Lambda(A)$ is positive definite, then for $0 < \alpha_N < 1$ and $N > N'$, we define

$$M_{\alpha_{N'-1}}(\zeta_{N'}, A^{-1}) = M(\zeta_{N'}, A^{-1})$$

$$M_{\alpha_N}(\zeta_{N+1}, A^{-1}) = (1 - \alpha_N) M_{\alpha_{N-1}}(\zeta_N, A^{-1}) + \alpha_N \phi(\underline{x}_{N+1}) A \phi'(\underline{x}_{N+1}), \quad N > N' \tag{F.4}$$

$$V_{\alpha_{N-1}}(\underline{x}, \zeta_N, A^{-1}) = \phi'(\underline{x}) M^{-1}_{\alpha_{N-1}}(\zeta_N, A^{-1}) \phi(\underline{x}), \quad \underline{x} \varepsilon \chi, \ N > N' \ .$$

Lemma F.1

(1) $\{M_{\alpha_N}(\zeta_{N+1}, A^{-1}),\ N > N' - 1\}$ is a sequence of positive definite matrices in $\Lambda(A)$.

(2) The elements of $M_{\alpha_N}(\zeta_{N+1}, A^{-1})$, $N > N' - 1$ are polynomials in elements of $\underline{b}(A)$, the r'-dimensional vector associated with A where $r' = r(r-1)/2$.

(3) $|M_{\alpha_N}(\zeta_{N+1}, A^{-1})|$, $N > N' - 1$ is a polynomial in the elements of $\underline{b}(A)$.

(4) The elements of $M^{-1}_{\alpha_N}(\zeta_{N+1}, A^{-1})$, $N > N' - 1$ is a rational function in the elements of $\underline{b}(A)$.

Proof

(1) We shall prove the result by the method of mathematical induction. For $N = N'$ we have

$$M_{\alpha_{N-1}}(\zeta_N, A^{-1}) = M(\zeta_{N'}, A^{-1}) \in \Lambda(A) \quad .$$

Observe that by definition $M(\zeta_N{}', A^{-1})$ is positive definite and belongs to $\Lambda(A)$. Suppose that $M_{\alpha_{N-1}}(\zeta_N, A^{-1})$ is positive definite and belongs to $\Lambda(A)$. We also have

$$M_{\alpha_N}(\zeta_{N+1}, A^{-1}) = (1 - \alpha_N)M_{\alpha_{N-1}}(\zeta_N, A^{-1}) + \alpha_N\phi(\underline{x}_{N+1})A\phi'(\underline{x}_{N+1}) \ ,$$

and since $M_{\alpha_{N-1}}(\zeta_N, A^{-1})$, $\phi(\underline{x}_{N+1})A\phi'(\underline{x}_{N+1}) \in \Lambda(A)$, the convexity of $\Lambda(A)$ implies that $M_{\alpha_N}(\zeta_{N+1}, A^{-1}) \in \Lambda(A)$. Next we need to

show that $M_{\alpha_N}(\zeta_{N+1}, A^{-1})$ is positive definite. This is done by means of the following result in Graybill (1969, p. 330):

If P is positive definite and Q is positive semidefinite then $|P - Q| < |P|$.

In our case, $P = M_{\alpha_N}(\zeta_{N+1}, A^{-1})$ and $Q = \alpha_N\phi(\underline{x}_{N+1})A\phi'(\underline{x}_{N+1})$ and we have

$$|M_{\alpha_N}(\zeta_{N+1}, A^{-1})| > |1 - \alpha_N|\; |M_{\alpha_{N-1}}(\zeta_N, A^{-1})| \quad .$$

Now since $M_{\alpha_{N-1}}(\zeta_N, A^{-1})$ is assumed to be positive definite and $\alpha_N < 1$ we have $|M_{\alpha_N}(\zeta_{N+1}, A^{-1})| > 0$. Since $M_{\alpha_N}(\zeta_{N+1}, A^{-1})$ is positive semidefinite we conclude that $M_{\alpha_N}(\zeta_{N+1}, A^{-1})$ must be positive definite.

(2) Let $N > N' - 1$. We have

$$M_{\alpha_N}(\zeta_{N+1}, A^{-1}) = \sum_{u=1}^{p'} \lambda_u^{N+1}\phi(\underline{x}_u^{N+1})A\phi'(\underline{x}_u^{N+1}) \tag{F.5}$$

where $0 < \lambda_u^{N+1} < 1$ with $\sum_{u=1}^{p'} \lambda_u^{N+1} = 1$.

Also

$$\phi(\underline{x}_u^{N+1})A\phi'(\underline{x}_u^{N+1})$$

$$=$$

$$\begin{bmatrix} \underline{f}_1(\underline{x}_u^{N+1}) & \underline{0} & \cdots & \underline{0} \\ \underline{0} & \underline{f}_2(\underline{x}_u^{N+1}) & \cdots & \underline{0} \\ \vdots & & \ddots & \vdots \\ 0 & & \cdots & \underline{f}_r(\underline{x}_u^{N+1}) \end{bmatrix} \begin{bmatrix} 1 & a_{12} & \cdots & a_{1r} \\ a_{21} & 1 & \cdots & a_{2r} \\ \\ a_{r1} & & \cdots & 1 \end{bmatrix} \begin{bmatrix} \underline{f}_1'(\underline{x}_u^{N+1}) & \underline{0}' & \cdots & \underline{0}' \\ \underline{0}' & \underline{f}_2'(\underline{x}_u^{N+1}) & & \underline{0}' \\ & & \ddots & \vdots \\ \underline{0}' & & \cdots & \underline{f}'_r(\underline{x}_u^{N+1}) \end{bmatrix}$$

$$=$$

$$\begin{bmatrix} \underline{f}_1(\underline{x}_u^{N+1})\underline{f}_1'(\underline{x}_u^{N+1}) & a_{12}\underline{f}_1(\underline{x}_u^{N+1})\underline{f}_2'(\underline{x}_u^{N+1}) & \cdots & a_{1r}\underline{f}_1(\underline{x}_u^{N+1})\underline{f}_r'(\underline{x}_u^{N+1}) \\ \\ a_{r1}f_r(\underline{x}_u^{N+1})f_1'(\underline{x}_u^{N+1}) & & \cdots & \underline{f}_r(\underline{x}_u^{N+1})\underline{f}_r'(\underline{x}_u^{N+1}) \end{bmatrix}.$$

It is clear from the above matrix equality that each element of $\phi(\underline{x}_u^{N+1})A\phi'(\underline{x}_u^{N+1})$ is a polynomial in $a_{12}, a_{13}, \ldots, a_{r-1,r}$ which are the elements of $\underline{b}(A)$. Now from (F.5) it is clear that each element of $M_{\alpha_N}\zeta_{N+1}, A^{-1})$ is a finite linear combination of the corresponding elements of $\phi(\underline{x}_u^{N+1})A\phi'(\underline{x}_u^{N+1})$, $u = 1, 2, \ldots, p'$ and hence is a polynomial in elements of $\underline{b}(A)$.

(3) Follows from (2) above and Theorem B.1(1).

(4) Follows from (2) above and Theorem B.1(2).

Lemma F.2

Let $\zeta\varepsilon\mathbf{H}$ amd $A\varepsilon\mathbf{A}'$. Then each element of $M(\zeta, A^{-1})$ is a polynomial function in $\underline{x}_1, \underline{x}_2, \ldots \underline{x}_{p'}, \lambda_1, \lambda_2, \ldots, \lambda_{p'}, a_{12}, a_{13}, \ldots, a_{r-1,r}$ defined on $\chi^{p'} \times \mathbf{U} \times [-1, 1]^{r'}$, where $r' = r(r-1)/2$, and $\mathbf{U} = \{\underline{\lambda}: \underline{\lambda} = (\lambda_1, \lambda_2, \ldots \lambda_{p'})$ and $0 < \lambda_u < 1$ with $\sum_{u=1}^{p'} \lambda_u = 1\}$.

Proof

By (F.3) we have

$$M(\zeta, A^{-1}) = \sum_{u=1}^{p'} \lambda_u \phi(\underline{x}_u) A \phi'(\underline{x}_u)$$

where $0 < \lambda_u < 1$ with $\sum_{u=1}^{p'} \lambda_u = 1$.

As shown in the proof of Lemma F.1(2) we have

$$\phi(\underline{x}_u)A\phi'(\underline{x}_u)$$

$$= \begin{bmatrix} \underline{f}_1(\underline{x}_u)\underline{f}_1'(\underline{x}_u) & a_{12}\underline{f}_1(\underline{x}_u)\underline{f}_2'(\underline{x}_u) & \cdots & a_{1r}\underline{f}_1(\underline{x}_u)\underline{f}_r'(\underline{x}_u) \\ \vdots & & & \vdots \\ a_{r1}\underline{f}_r(\underline{x}_u)\underline{f}_1'(\underline{x}_u) & & \cdots & \underline{f}_r(\underline{x}_u)\underline{f}_r'(\underline{x}_u) \end{bmatrix}.$$

Therefore

$M(\zeta, A^{-1}) =$

$$\begin{bmatrix} \sum_{u=1}^{p'} \lambda_u f_1(\underline{x}_u) f_1'(\underline{x}_u) & a_{12} \sum_{u=1}^{p'} \lambda_u \underline{f}_1(\underline{x}_u) \underline{f}_2'(\underline{x}_u) & \cdots & a_{1r} \sum_{u=1}^{p'} \lambda_u \underline{f}_1(x_u) \underline{f}_r'(\underline{x}_u) \\ a_{r1} \sum_{u=1}^{p'} \lambda_u \underline{f}_r(\underline{x}_u) \underline{f}_1'(\underline{x}_u) & & \cdots & \sum_{u=1}^{p'} \lambda_u \underline{f}_r(\underline{x}_u) \underline{f}_r'(\underline{x}_u) \end{bmatrix} .$$

Now $f_i(\underline{x},)$, $i = 1, 2, \ldots, r$ is a polynomial function in $\underline{x}$ and hence it is clear from the above equality that each element of $M(\zeta, A^{-1})$ is a polynomial function in $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_{p'}, \lambda_1, \lambda_2, \ldots, \lambda_{p'}, a_{12}, a_{13}, \ldots, a_{r-1,r}$ defined on $\chi^{p'} \times \mathbf{U} \times [-1, 1]^{r'}$.

Lemma F.3

Let $\zeta \varepsilon H$ be such that $M(\zeta, A_0^{-1})$ is positive definite for some $A_0 \varepsilon \mathbf{A}'$. Then $M(\zeta, A^{-1})$ is positive definite for all $A \varepsilon \mathbf{A}'$.

Proof.

$$M(\zeta, A_0^{-1}) = \sum_{u=1}^{p'} \lambda_u \phi(\underline{x}_u) A_0 \phi'(\underline{x}_u)$$

where $0 < \lambda_u < 1$ with $\sum_{u=1}^{p'} \lambda_u = 1$.

We can also prove that

$$\sum_{u=1}^{p'} \lambda_u \phi(\underline{x}_u) A \phi'(\underline{x}_u) = \tilde{F}' (A \otimes I_{p'}) \tilde{F} \quad ,$$

where $\tilde{F}$ is a rp'xp matrix such that

$$\tilde{F} = \begin{bmatrix} \tilde{F}_1 & 0 & \dots & 0 \\ & \tilde{F}_2 & & \\ & & \ddots & \\ 0 & & & \tilde{F}_r \end{bmatrix} ,$$

in which $\tilde{F}_i$ is a $p'xp_i$ matrix of full column rank whose u^{th} row contains the elements of $\lambda_u^{1/2} \underline{f}_i'(\underline{x}_u)$, $u = 1, 2, \dots, p'$. Thus

$$M(\zeta, A_0^{-1}) = \tilde{F}'(A \otimes I) \tilde{F}$$

$$= \tilde{F}'(A_0^{1/2} \otimes I_{p'})(A_0^{1/2} X I_{p'}) \tilde{F}$$

($A_0^{1/2}$ exists since A_0 is positive definite)

Also if $A \varepsilon \mathbf{A}'$ then $A^{1/2}$ exists (since A is positive definite) and

$$\text{rank}[(A^{1/2} \otimes I_{p'})\tilde{F}] = \text{rank}[(A^{1/2} \otimes I_{p'})(A_0^{-1/2} \otimes I_{p'})(A_0^{1/2} \otimes I_{p'})\tilde{F}]$$

$$= \text{rank}[(A_0^{1/2} \otimes I_{p'})\tilde{F}]$$

$$= \text{rank}[\tilde{F}'(A_0 \otimes I_{p'})\tilde{F}]$$

$$= \text{rank}[M(\zeta,\ A_0^{-1})]$$

$$= p \quad \text{(since, } M(\zeta,\ A_0^{-1}) \text{ is positive definite).}$$

It follows that

$$\text{rank}[M(\zeta,\ A^{-1})] = \text{rank}[\tilde{F}'(A \otimes I_{p'})\tilde{F}]$$

$$= \text{rank}[(A^{1/2} \otimes I_{p'})\tilde{F}]$$

$$= p \quad .$$

Now since $M(\zeta, A^{-1})$ is of the form $\tilde{F}'(A^{1/2}\otimes I)(A^{1/2}\otimes I)\tilde{F}$, it is positive semidefinite and therefore positive definite.

Lemma F.4

Let $A_0 \in \mathbf{A}'$ and ζ_N' be a discrete design such that $M(\zeta_{N'}, A_0^{-1})$ is positive definite and let $\{A_N, N > N'\}$ be any sequence in $\mathbf{A}'$ such that $\underline{b}(A_N)$ converges to $\underline{b}(A_0)$. Then for any α_N such that $0 < \alpha_N < 1$, there exists a positive integer N_1 such that

$$\log|M_{\alpha_N}(\zeta_{N+1}, A_N^{-1})| > \log|M_{\alpha_{N-1}}(\zeta_N, A_{N-1}^{-1})| +$$

$$\alpha_N\{tr[A_N V_{\alpha_{N-1}}(\underline{x}_{N+1}, \zeta_N, A_N^{-1})]$$

$$-p\} + r_N - 2r'pe_{min}^{-1}(A_0)(e_N^o + e_{N-1}^o), \; N > N_1$$

where $r_N = O(\alpha_N)$, e_N^o denotes the r'-dimensional Euclidean distance between $\underline{b}(A_N)$ and $\underline{b}(A_0)$, and e_{min} denotes the smallest eigenvalue of the matrix inside parantheses.

Proof

We shall first prove that $M_{\alpha_{N-1}}(\zeta_N, A^{-1})$ is positive definite for any $\underline{b}(A) \in B(A_0, \delta)$ and $N > N_1$ where $B(A_0, \delta)$ is a spherical

neighborhood of radius δ centered at $\underline{b}(A_0)$, and N_1 is a positive integer. Now A_0 is positive definite implies $e_{min}(A_0) > 0$, and since $e_{min}(A)$ is continuous in the elements of A there exists $\delta > 0$ such that $e_{min}(A) > e_{min}(A_0)/2 > 0$ (hence A is positive definite) for all A, such that $\underline{b}(A) \varepsilon B(A_0, \delta)$. Therefore, $A \varepsilon \mathbf{A}'$, whenever $\underline{b}(A) \varepsilon B(A_0, \delta)$. Also since $\underline{b}(A_N)$ converges to $\underline{b}(A_0)$ there exists $N_1 > N'$ such that $\underline{b}(A_N) \varepsilon B(A_0, \delta)$ for all $N > N_1$. Now since $M(\zeta_{N'}, A_0^{-1})$ is positive definite by hypothesis, it follows that $M_{\alpha_{N-1}}(\zeta_N, A_0^{-1})$, $N > N_1$ is positive definite in $\Lambda(A_0)$ in view of Lemma F.1(1). Moreover, since $\underline{b}(A) \varepsilon B(A_0, \delta)$ implies $A\varepsilon\mathbf{A}'$ we have from Lemma F.3 that $M_{\alpha_{N-1}}(\zeta_N, A^{-1})$, $N > N_1$ is positive definite whenever $\underline{b}(A)\varepsilon B(A_0, \delta)$. This implies that $\log|M_{\alpha_{N-1}}(\zeta_N, A^{-1})|$ is well defined whenever $\underline{b}(A) \varepsilon B(A_0, \delta)$ and $N > N_1$. Let $N > N_1 + 1$ be fixed. Then $\underline{b}(A_N)$, $\underline{b}(A_{N-1}) \varepsilon B(A_0, \delta)$ and for fixed α_{N-1}, $\log|M_{\alpha_{N-1}}(\zeta_N, A^{-1})|$ is a polynomial function of $\underline{b}(A)$ defined on $B(A_0, \delta)$. Hence we denote $\log|M_{\alpha_{N-1}}(\zeta_N, A^{-1})|$ by $h(A)$. Now $h(A)$ is differentiable on $B(A_0, \delta)$ and therefore we use the Mean Value Theorem in many variables (see Gillespie, 1951, p. 60) to write

$$h(A) = h(A_0) + \sum_{\ell=1}^{r'} (b_\ell - b_\ell^0)\left[\frac{\partial}{\partial b_\ell} h(A)\right]_{A=A(\theta)}, \quad \text{where } 0 < \theta < 1,$$

$$\underline{b}(A) \varepsilon B(A_0, \delta) \quad \text{with} \quad \underline{b}(A_0) = (b_1^0, b_2^0, \ldots, b_{r'}^0)' \tag{F.6}$$

$$\text{and } \underline{b}(A(\theta)) = (1 - \theta)\underline{b}(A) + \theta\underline{b}(A_0) \quad .$$

Notice that $\underline{b}(A(\theta)) \in B(A_0, \delta)$. Next we prove that

$|\frac{\partial}{\partial b_\ell} h(A)| < 2pe_{min}^{-1}(A_0)$ whenever $\underline{b}(A) \in B(A_0, \delta)$, and $1 \leq l \leq r'$. Now

$$\frac{\partial}{\partial b_\ell} h(A) = \frac{\partial}{\partial b_\ell} \log|M_{\alpha_{N-1}}(\zeta_N, A^{-1})|$$

$$= tr[M^{-1}_{\alpha_{N-1}}(\zeta_N, A^{-1})\frac{\partial}{\partial b_\ell}\{M_{\alpha_{N-1}}(\zeta_N, A^{-1})\}] \qquad (F.7)$$

(using equation 1.134 on p. 21 in Fedorov (1972)).

Also from Lemma F.1(1) we have

$$M_{\alpha_{N-1}}(\zeta_N, A^{-1}) = \sum_{u=1}^{p'} \lambda_u^N \phi(\underline{x}_u^N) A \phi'(\underline{x}_u^N) \quad \text{with } 0 \leq \lambda_u^N \leq 1 \text{ and}$$

$$\sum_{u=1}^{p'} \lambda_u^N = 1 \quad .$$

Therefore

$$\frac{\partial}{\partial b_\ell} M_{\alpha_{N-1}}(\zeta_N, A^{-1})$$

$$= \sum_{u=1}^{p'} \lambda_u^N \frac{\partial}{\partial b_\ell} \phi(\underline{x}_u^N) A \phi'(\underline{x}_u^N)$$

$$= \sum_{u=1}^{p'} \lambda_u^N \frac{\partial}{\partial b_\ell} \begin{bmatrix} \underline{f}_1(\underline{x}_u^N)\underline{f}_1'(\underline{x}_u^N) & a_{12}\underline{f}_1(\underline{x}_u^N)\underline{f}_2'(\underline{x}_u^N) & \cdots & a_{1r}\underline{f}_1(\underline{x}_u^N)\underline{f}_r'(\underline{x}_u^N) \\ \\ a_{r1}\underline{f}_r(\underline{x}_u^N)\underline{f}_1'(\underline{x}_u^N) & & \cdots & \underline{f}_r(\underline{x}_u^N)\underline{f}_r'(\underline{x}_u^N) \end{bmatrix}$$

, where $A=[a_{ij}]$

$$= \sum_{u=1}^{p'} \lambda_u^N \phi(\underline{x}_u^N) B^\ell \phi'(\underline{x}_u)$$

where B^ℓ is the rxr symmetric matrix such that its ij^{th} and ji^{th} elements are equal to 1 and all other elements are equal to zero, since b_ℓ is equal to some a_{ij}. Substituting for $\frac{\partial}{\partial b_\ell} M_{\alpha_{N-1}}(\zeta_N, A^{-1})$ in (F.7) we obtain

$$\frac{\partial}{\partial b_\ell} h(A) = tr[M^{-1}_{\alpha_{N-1}}(\zeta_N, A^{-1}) \frac{\partial}{\partial b_\ell}\{M_{\alpha_{N-1}}(\zeta_N, A^{-1})\}]$$

$$= tr[M^{-1}_{\alpha_{N-1}}(\zeta_N, A^{-1})\{\sum_{u=1}^{p'} \lambda_u^N \phi(\underline{x}_u^N) B^\ell \phi'(\underline{x}_u^N)\}]$$

$$= tr[M_1 M_1'\{\sum_{u=1}^{p'} \lambda_u^N \phi(\underline{x}_u^N) B^\ell \phi'(\underline{x}_u^N)\}] \text{, where } M^{-1}_{\alpha_{N-1}}(\zeta_N, A^{-1}) = M_1 M_1'$$

$$= \sum_{u=1}^{p'} \lambda_u tr[M_1'\phi(\underline{x}_u^N)B^{\ell}\phi'(\underline{x}_u^N)M_1] \quad .$$

Hence

$$\left|\frac{\partial}{\partial b_{\ell}} h(A)\right| < \left| \sum_{u=1}^{p'} \lambda_u^N tr[M_1'\phi(\underline{x}_u^N)B^{\ell}\phi'(\underline{x}_u^N)M_1]\right|$$

$$< \sum_{u=1}^{p'} \lambda_u^N |tr[M_1'\phi(\underline{x}_u^N)B^{\ell}\phi'(\underline{x}_u^N)M_1]| \quad . \qquad (F.9)$$

Now

$$e_{min}(B^{\ell})tr[M_1'\phi(\underline{x}_u^N)\phi'(\underline{x}_u^N)M_1]$$

$$< tr[M_1'\phi(\underline{x}_u^N)B^{\ell}\phi'(\underline{x}_u^N)M_1]$$

$$< e_{max}(B^{\ell})tr[M_1'\phi(\underline{x}_u^N)\phi'(\underline{x}_u^N)M_1] \quad , \quad 1 < u < p'$$

(using Theorem B.3).

Thus

$$-tr[M_1'\phi(\underline{x}_u^N)\phi'(\underline{x}_u^N)M_1] < tr[M_1'\phi(\underline{x}_u^N)B^{\ell}\phi'(\underline{x}_u^N)M_1] < tr[M_1'\phi(x_u^N)\phi'(\underline{x}_u^N)M_1]$$

(since $e_{min}(B^{\ell}) = -1$ and $e_{max}(B^{\ell}) = 1$ from Theorem B.2),

and so

$$|tr[M_1'\phi(\underline{x}_u^N)B^{\ell}\phi'(\underline{x}_u^N)M_1]| \le tr[M_1'\phi(\underline{x}_u^N)\phi'(\underline{x}_u^N)M_1] \quad . \tag{F.10}$$

Moreover

$$e_{min}(A)tr[M_1'\phi(\underline{x}_u^N)\phi'(\underline{x}_u^N)M_1] \le tr[M_1'\phi(\underline{x}_u^N)A\phi'(\underline{x}_u^N)M_1]$$

(using Theorem B.3).

It follows that

$$tr[M_1'\phi(\underline{x}_u^N)\phi'(\underline{x}_u^N)M_1] \le e_{min}^{-1}(A)tr[M_1'\phi(\underline{x}_u^N)A\phi'(\underline{x}_u^N)M_1]$$

$$\le 2e_{min}^{-1}(A_0)tr[M_1'\phi(\underline{x}_u^N)A\phi'(\underline{x}_u^N)M_1]$$

(recall that $e_{min}(A) \ge e_{min}(A_0)/2 > 0$, whenever $\underline{b}(A) \in B(A_0, \delta)$).

Now using inequality (F.10) we get

$$|tr[M_1'\phi(\underline{x}_u^N)B^{\ell}\phi'(\underline{x}_u^N)M_1]| \le 2e_{min}^{-1}(A_0)tr[M_1'\phi(\underline{x}_u^N)A\phi'(\underline{x}_u^N)M_1] \quad .$$

Finally using the above inequality together with (F.9) we obtain

$$\left|\frac{\partial}{\partial b_\ell} h(A)\right|$$

$$< \sum_{u=1}^{p'} \lambda_u^N \left|tr[M_1'\phi(\underline{x}_u^N)B^\ell\phi'(\underline{x}_u^N)M_1]\right|$$

$$= \sum_{u=1}^{p'} 2\lambda_u^N e_{min}^{-1}(A_0)tr[M_1'\phi(\underline{x}_u^N)A\phi'(\underline{x}_u^N)M_1]$$

$$= 2e_{min}^{-1}(A_0)\sum_{u=1}^{p'} \lambda_u^N tr[M_1M_1'\phi(\underline{x}_u^N)A\phi'(\underline{x}_u^N)]$$

$$= 2e_{min}^{-1}(A_0)\sum_{u=1}^{p'} \lambda_u^N tr[M_{\alpha_{N-1}}^{-1}(\zeta_N, A^{-1})\phi(\underline{x}_u^N)A\phi'(\underline{x}_u^N)]$$

$$= 2e_{min}^{-1}(A_0)tr[M_{\alpha_{N-1}}^{-1}(\zeta_N, A^{-1})\sum_{u=1}^{p'} \lambda_u^N\phi(\underline{x}_u^N)A\phi'(\underline{x}_u^N)]$$

$$= 2e_{min}^{-1}(A_0)tr[M_{\alpha_{N-1}}^{-1}(\zeta_N, A^{-1})M_{\alpha_{N-1}}(\zeta_N, A^{-1})]$$

$$= 2e_{min}^{-1}(A_0)p \quad .$$

Consequently

$$\left|\frac{\partial}{\partial b_\ell} h(A)\right| < 2e_{min}^{-1}(A_0)p \text{ whenever } \underline{b}(A) \in B(A_0, \delta), \text{ and } 1 < \ell < r' .$$

In particular, $\left|\left[\frac{\partial}{\partial b_\ell} h(A)\right]_{A=A(\theta)}\right| < 2e_{min}^{-1}(A_o)p$,since $\underline{b}(A(\theta)) \in B(A, \delta)$.
Now from (F.6) we have

$$h(A) = h(A_0) + \sum_{\ell=1}^{r'} (b_\ell - b_\ell^0) \frac{\partial}{\partial b_\ell} h(A(\theta))$$

$$\text{where } \left[\frac{\partial}{\partial b_\ell} h(A)\right]_{A=A(\theta)} = \frac{\partial}{\partial b_\ell} h(A(\theta)) .$$

Thus

$$|h(A) - h(A_0)| = \left|\sum_{\ell=1}^{r'} (b_\ell - b_\ell^0) \frac{\partial}{\partial b_\ell} h(A(\theta))\right|$$

$$< \sum_{\ell=1}^{r'} |b_\ell - b_\ell^0| \left|\frac{\partial}{\partial b_\ell} h(A(\theta))\right|$$

$$< \sum_{\ell=1}^{r'} 2e_{min}^{-1}(A_o)p\, e^0 \text{ where } e^0 \text{ is the}$$

Euclidean distance between $\underline{b}(A_0)$ and $\underline{b}(A)$.

$$= 2pe^0 e_{min}^{-1}(A_0) r' \quad ,$$

i.e.,

$|h(A) - h(A_0)| < 2pe^0 e_{min}^{-1}(A_0)r'$ whenever $\underline{b}(A) \varepsilon B(A_0, \delta)$.

Recall that $\underline{b}(A_N)$, $\underline{b}(A_{N-1}) \varepsilon B(A_0, \delta)$. Therefore, replacing A by A_N and A_{N-1} we get

$$|h(A_N) - h(A_0)| < 2pe_N^0 e_{min}^{-1}(A_0)r' \quad ,$$

and

$$|h(A_{N-1}) - h(A_0)| < 2pe_{N-1}^0 e_{min}^{-1}(A_0)r' \quad .$$

These two inequalitites imply that

$$-2pr'e_N^0 e_{min}^{-1}(A_0) < h(A_N) - h(A_0) < 2pr'e_N^0 e_{min}^{-1}(A_0) \quad ,$$

and

$$-2pr'e_{N-1}^0 e_{min}^{-1}(A_0) < h(A_{N-1}) - h(A_0) < 2pr'e_{N-1}^0 e_{min}^{-1}(A_0)$$

from which we obtain

$$h(A_N) - h(A_{N-1}) > -2pr'e_{min}^{-1}(A_0)(e_N^0 + e_{N-1}^0) \quad .$$

Substituting for $h(A_N)$ and $h(A_{N-1})$, we finally have

$$\log|M_{\alpha_{N-1}}(\zeta_N, A_N^{-1})| > \log|M_{\alpha_{N-1}}(\zeta_N, A_{N-1}^{-1})| - 2p r' e_{min}^{-1}(A_0)(e_N^0 + e_{N-1}^0),$$

$$N > N_1 + 1 \quad . \qquad (F.11)$$

The rest of the proof somewhat parallels the arguments given in Silvey (1980, p. 36):

The matrix $M_{\alpha_N}(\zeta_{N+1}, A_N^{-1})$ was defined as

$$M_{\alpha_N}(\zeta_{N+1}, A_N^{-1}) = (1 - \alpha_N)M_{\alpha_{N-1}}(\zeta_N, A_N^{-1}) + \alpha_N \phi(\underline{x}_{N+1}) A_N \phi'(\underline{x}_{N+1}).$$

For given $N > N_1 + 1$, $\log|M_{\alpha_N}(\zeta_{N+1}, A_N^{-1})|$ is a differentiable function of α_N and therefore we can use Taylor's Theorem about $\alpha_N = 0$ to get

$$\log|M_{\alpha_N}(\zeta_{N+1}, A_N^{-1})| = \log|M_{\alpha_{N-1}}(\zeta_N, A_N^{-1})| +$$

$$\alpha_N \left[\frac{\partial}{\partial \alpha_N} \log|M_{\alpha_N}(\zeta_{N+1}, A_N^{-1})|\right]_{\alpha_N=0} + r_N$$

where $r_N = 0(\alpha_N)$

$$= \log|M_{\alpha_{N-1}}(\zeta_N, A_N^{-1})| + \alpha_N \mathrm{tr}[M_{\alpha_{N-1}}^{-1}(\zeta_N, A_N^{-1})\{\phi(\underline{x}_{N+1})A_N\phi'(\underline{x}_{N+1}) - M_{\alpha_{N-1}}(\zeta_N, A_N^{-1})\}] + r_N$$

(using equation 1.1.34 on p. 21 in Fedorov (1972))

$$= \log|M_{\alpha_{N-1}}(\zeta_N, A_N^{-1})| + \alpha_N\{\mathrm{tr}[M_{\alpha_{N-1}}^{-1}(\zeta_N, A_N^{-1})\phi(\underline{x}_{N+1})A_N\phi'(\underline{x}_{N+1})]-p\} + r_N$$

$$= \log|M_{\alpha_{N-1}}(\zeta_N, A_N^{-1})| + \alpha_N\{\mathrm{tr}\{A_N\phi'(\underline{x}_{N+1})M_{\alpha_{N-1}}^{-1}(\zeta_N, A_N^{-1})\phi(\underline{x}_{N+1})]-p\} + r_N$$

$$= \log|M_{\alpha_{N-1}}(\zeta_N, A_N^{-1})| + \alpha_N\{\mathrm{tr}A_N V_{\alpha_{N-1}}(\underline{x}_{N+1}, \zeta_N, A_N^{-1})] - p\} + r_N$$

$$> \log|M_{\alpha_{N-1}}(\zeta_N, A_{N-1}^{-1})| - 2pr'(e_N^0 + e_{N-1}^0)e_{min}^{-1}(A_0) + \alpha_N\{\mathrm{tr}[A_N V_{\alpha_{N-1}}(\underline{x}_{N+1}, \zeta_N, A_N^{-1})] - p\} + r_N, \quad \text{(using inequality F.11).}$$

We are now ready to prove the main result.

Theorem 2.4

Suppose that $\sum_{u=1}^{N} \hat{e}_u$ converges in probability to some random variable e, then for a given $\delta > 0$, there exists an integer $N > 0$ such that

$$\sup_{\underline{x}\varepsilon X} tr[\hat{A}_N V(\underline{x}, \hat{\zeta}_N, \hat{A}_N^{-1})] - p < \delta$$

with probability 1 (see (2.20) for the definition of $\hat{A}_N$; Section 2.6 for the definitions of $\hat{\zeta}_N$, $\hat{e}_N$).

Proof

Since $\hat{e}_u \geq 0$, $\sum_{u=1}^{N} \hat{e}_u$ is a nondecreasing sequence of random variables. Thus $\{\sum_{u=1}^{N} \hat{e}_u, N \geq 1\}$ is a nondecreasing random sequence which converges in probability to a random variable e, and hence $\sum_{u=1}^{N} \hat{e}_u$ converges w.p.1 to the same random variable e (Billingsley, 1979, p. 236), i.e., if $\Omega_0 = \{\omega: \lim_{N\to\infty} \sum_{u=1}^{N} \hat{e}_u(\omega) = e(\omega)\}$, then $P(\Omega_0) = 1$. Let $\omega \varepsilon \Omega_0$. Then $\lim_{N\to\infty} \hat{e}_N(\omega) = 0$. Equivalently, $\lim_{N\to\infty} \underline{b}(\hat{A}_N(\omega)) = \underline{b}(A)$. Using Lemma F.4 with $A_N = \hat{A}_N(\omega)$, $\alpha_N = 1/(N+1)$, $\zeta_N = \hat{\zeta}_N(\omega)$, and $A_0 = A$ (as defined in (2.19)), we have

$$\log|M(\hat{\zeta}_{N+1}(\omega),\ \hat{A}_N^{-1}(\omega))|$$

$$> \log|M(\hat{\zeta}_N(\omega),\ \hat{A}^{-1}{}_{N-1}(\omega))| + (1/N+1)\{tr[\hat{A}_N(\omega)$$

$$V(\hat{\underline{x}}_{N+1}(\omega),\ \hat{\zeta}_N(\omega),\ \hat{A}_N^{-1}(\omega))] - p\}$$

$$+ r_N - 2r'p\ e_{min}^{-1}(A)(\hat{e}_N(\omega) + \hat{e}_{N-1}(\omega)) \quad , \quad N > N_1(\omega)$$

(note that when $\alpha_N = 1/(N+1)$, $M_{\alpha_N}(\zeta_N, A^{-1})$ and $V_{\alpha_N}(\underline{x},\ \zeta_N,\ A^{-1})$ reduce to $M(\zeta_N,\ A^{-1})$ and $V(\underline{x},\ \zeta_N,\ A^{-1})$ mentioned in Sections 2.3-2.6)

$$= \log|M(\hat{\zeta}_N(\omega),\ \hat{A}_{N-1}^{-1}(\omega))| + (1/N+1)[\sup_{\underline{x}\varepsilon X} tr\{\hat{A}_N(\omega)V(\underline{x},\hat{\zeta}_N(\omega),\hat{A}_N^{-1}(\omega))\} - p]$$

$$+ r_N - 2r'p\ e_{min}^{-1}(A)(\hat{e}_N(\omega) + \hat{e}_{N-1}(\omega)) \quad , \quad N > N_1(\omega) \quad .$$

Recall that $tr[\hat{A}_N V(\hat{\underline{x}}_{N+1},\ \hat{\zeta}_N,\ \hat{A}_N^{-1})] = \sup_{\underline{x}\varepsilon X} tr[\hat{A}_N V(\underline{x},\ \hat{\zeta}_N,\ \hat{A}_N^{-1})]$

(see Section 2.6).

The proof from this point on, closely follows the arguments given in Silvey (1980, pp. 35-36). Suppose

$$\sup_{\underline{x}\epsilon X} tr[\hat{A}_N(\omega)V(\underline{x}, \hat{\zeta}_N(\omega), \hat{A}_N^{-1}(\omega))] - p > \delta \text{ for all } N \quad .$$

Then

$$\log|M(\hat{\zeta}_{N+1}(\omega), \hat{A}_N^{-1}(\omega))|$$

$$> \log|M(\hat{\zeta}_N(\omega), \hat{A}_{N-1}^{-1}(\omega))| + \delta/(N+1) + r_N - 2r'pe_{min}^{-1}(A)\ (\hat{e}_N(\omega) + \hat{e}_{N-1}(\omega)),$$

$$N > N_1(\omega) \quad .$$

Using the above recursion formula, and by repeated substitution we have

$$\log|M(\hat{\zeta}_{N+1}(\omega)), \hat{A}_N^{-1}(\omega)| - \log|M(\hat{\zeta}_{N_1(\omega)+1}(\omega), \hat{A}_{N_1(\omega)}^{-1}(\omega))|$$

$$> -2r'pe_{min}^{-1}(A)\left(\sum_{u=N_1(\omega)+1}^{N} \hat{e}_u(\omega) + \sum_{u=N_1(\omega)}^{N} \hat{e}_u(\omega)\right) +$$

$$\delta \sum_{u=N_1(\omega)+1}^{N} 1/(u+1) + \sum_{u=N_1(\omega)+1}^{N} r_u \quad \text{where} \quad r_u = O(1/u+1)$$

$$= -2r'pe_{min}^{-1}(A)\left(\sum_{u=N_1(\omega)+1}^{N} \hat{e}_u(\omega) + \sum_{u=N_1(\omega)}^{N} \hat{e}_u(\omega)\right)$$

$$+ \sum_{u=N_1(\omega)+1}^{N} (1/u+1)\{\delta + (\sum_{u=N_1(\omega)+1}^{N} r_u / \sum_{u=N_1(\omega)+1}^{N} (1/u+1))\} \quad .$$

The limit of the second term on the right hand side as $N \to \infty$ is ∞ (see Silvey, 1980, p. 36). Therefore, the limit of the right hand side as $N \to \infty$ is $-2r'pe_{min}^{-1}(A)\{2e(\omega) - 2 \sum_{u=1}^{N_1(\omega)-1} \hat{e}_u - \hat{e}_{N_1(\omega)}\} + \infty = \infty$, implying that $\log|M(\hat{\zeta}_{N+1}(\omega), \hat{A}_N^{-1}(\omega))|$ diverges to ∞ for any $\omega \varepsilon \Omega_0$. Equivalently $\log|M(\zeta_{N+1}, \hat{A}_N^{-1})|$ diverges to ∞ w.p.1. But $\log|M(\zeta, A^{-1})|$, $\zeta \varepsilon \mathbf{H}$, $A \varepsilon \mathbf{A}'$ is a continuous function on the closed and bounded subspace $X^{p'} \times \mathbf{U} \times [-1, 1]^{r'}$. Hence $\log|M(\hat{\zeta}_{N+1}, \hat{A}_N^{-1})|$ is a bounded random function. This contradicts the fact that $\log|M(\hat{\zeta}_{N+1}, \hat{A}_N^{-1})|$ diverges to ∞ w.p.1. This contradiction establishes the fact that $\sup_{\underline{x} \varepsilon X} tr[\hat{A}_N V(\underline{x}, \hat{\zeta}_N, \hat{A}_N^{-1}] - p < \delta$ for some integer N w.p.1.

APPENDIX G
AN EXPRESSION FOR $F_{\Lambda_2'}$ $[M(\zeta), \underline{h}(\underline{x})\underline{h}'(\underline{x})]$

Only an outline of the proof is given and routine algebraic manipulations are omitted.

By definition 3.4

$$F_{\Lambda_2'} (M, \underline{h}\ \underline{h}') = \lim_{\varepsilon \to 0^+} (1/\varepsilon) [\Lambda_2'(\bar{M}) - \Lambda_2'(M)]$$

where $\bar{M} = (1 - \varepsilon)M + \varepsilon\underline{h}\ \underline{h}'$.

Recall that

$$\Lambda_2'(M) = tr[T^{-1}L\{I_r \otimes (M_{ZZ} - M_{ZX}M_{XX}^{-1}M_{XZ})\}L'] \quad .$$

Therefore

$$F_{\Lambda_2'}(M, \underline{h}\ \underline{h}') = \lim_{\varepsilon \to 0^+} tr\{T^{-1}L(I_r \otimes E)L'\}$$

where $E = (1/\varepsilon)\{\bar{M}_{ZZ} - \bar{M}_{ZX}\bar{M}_{XX}^{-1}\bar{M}_{XZ} - M_{ZZ} + M_{ZX}M_{XX}^{-1}M_{XZ}\}$

and

$$\bar{M} = \begin{bmatrix} \tilde{M}_{XX} & \tilde{M}_{XZ} \\ \tilde{M}_{ZX} & \tilde{M}_{ZZ} \end{bmatrix} .$$

Using the following identity from Dykstra (1971) for a nonsingular matrix A:

$$(A + \underline{x}_0\underline{x}_0')^{-1} = A^{-1} - \frac{A^{-1}\underline{x}_0\underline{x}_0'A^{-1}}{1+\underline{x}_0'A^{-1}\underline{x}_0}$$

with

$$A = (1 - \varepsilon)M_{XX} \quad \text{and} \quad \underline{x}_0 = \varepsilon^{1/2}\underline{a}$$

we have

$$\bar{M}_{XX}^{-1} = [(1 - \varepsilon)M_{XX} + \varepsilon\underline{a}\underline{a}']^{-1}$$

$$= (1 - \varepsilon)^{-1}M_{XX}^{-1} - \{\varepsilon c/(1 - \varepsilon)^2\}M_{XX}^{-1}\underline{a}\underline{a}'M_{XX}^{-1}$$

where $c^{-1} = 1 + \{\varepsilon/(1 - \varepsilon)\}\underline{a}'M_{XX}^{-1}\underline{a}$

i.e.,

$$\tilde{M}_{XX}^{-1} = (1 - \varepsilon)^{-1} M_{XX}^{-1} - \{\varepsilon/t(1 - \varepsilon)\}P$$

where $t = 1 - \varepsilon + \varepsilon \underline{a}' M_{XX}^{-1} \underline{a}$ and $P = M_{XX}^{-1} \underline{a}\,\underline{a}' M_{XX}^{-1}$.

With the use of the above equality and cancellation of terms we get

$$E = -M_{ZZ} + M_{ZX} M_{XX}^{-1} M_{XZ} + \underline{bb}' - M_{ZX} M_{XX}^{-1} \underline{ab}' - \underline{ba}' M_{XX}^{-1} M_{XZ}$$

$$-\{\varepsilon/(1 - \varepsilon)\}\underline{ba}' M_{XX}^{-1} \underline{ab}' + \{(1 - e)/t\} M_{ZX} P M_{XZ}$$

$$+ (\varepsilon/t) M_{ZX} P \underline{ab}' + (\varepsilon/t) \underline{ba}' P M_{XZ} + \{\varepsilon^2/(1 - \varepsilon)t\} \underline{ba}' P \underline{ab}' \quad .$$

Using the above equality and evaluating the limit of each term in $tr[T^{-1} L(I_r \otimes E) L']$ as ε goes to 0^+ we obtain

$$F_{\Lambda_2'}[M(\zeta), \underline{h}(\underline{x})\underline{h}'(\underline{x})] = tr[T^{-1} L\{I_r \otimes ((\underline{b}(\underline{x}) - \underline{v}(\underline{x}, \zeta))$$

$$(b'(\underline{x}) - \underline{v}'(x, \zeta)))\} L'] - \Lambda_2'[M(\zeta)]$$

where $\underline{v}(\underline{x}, \zeta) = M_{ZX}(\zeta) M_{XX}^{-1}(\zeta) \underline{a}(\underline{x})$.

APPENDIX H
PROOF OF LEMMA 3.1

Lemma 3.1

Λ_2' is concave on $\mathbf{M}$, differentiable on $\mathbf{M}'$, and a Λ_2'-optimal measure exists.

Proof

Recall that

$$\mathbf{M} = \{M(\zeta),\ \zeta \in H\}, \text{ and } \mathbf{M}' = \{M(\zeta): M_{XX}(\zeta) \text{ is nonsingular}\} \quad .$$

Suppose that $\underline{\omega} = (\lambda_1, \lambda_2, \ldots, \lambda_{m'}, \underline{x}_1, \underline{x}_2, \ldots, \underline{x}_{m'})'$ is the point in $\mathbf{U} \times \chi^{m'}$ which is associated with $M(\zeta) \in \mathbf{M}$. Then

$$M(\zeta) = \begin{bmatrix} M_{XX}(\zeta) & M_{XZ}(\zeta) \\ & \\ M_{ZX}(\zeta) & M_{ZZ}(\zeta) \end{bmatrix}$$

where $M_{XX}(\zeta) = \sum_{u=1}^{m'} \lambda_u \underline{a}(\underline{x}_u)\underline{a}'(\underline{x}_u)$, $M_{XZ}(\zeta) = \sum_{u=1}^{m'} \lambda_u \underline{a}(\underline{x}_u)\underline{b}'(\underline{x}_u)$, $M_{ZX}(\zeta) = \sum_{u=1}^{m'} \lambda_u \underline{b}(\underline{x}_u)\underline{a}'(\underline{x}_u)$, and $M_{ZZ}(\zeta) = \sum_{u=1}^{m'} \lambda_u \underline{b}(\underline{x}_u)\underline{b}'(\underline{x}_u)$.

Also

$$\Lambda_2'[M(\zeta)] = \operatorname{tr}[T^{-1}L\{I_r \otimes (M_{ZZ}(\zeta) - M_{ZX}(\zeta)M_{XX}^{-1}(\zeta)M_{XZ}(\zeta))\}L'] \quad .$$

It is clear that the elements of $M_{XX}(\zeta)$, $M_{XZ}(\zeta)$, $M_{ZX}(\zeta)$, and $M_{ZZ}(\zeta)$ are polynomial functions of $\underline{\omega}$ and hence $\Lambda_2'[M(\zeta)]$ is a continuous function on the closed and bounded subspace $U \times X^{m'}$. Hence a Λ_2'-optimal measure exists, since a continuous function on a closed and bounded subspace attains its optimum value.

Next, the differentiability of Λ_2' on **M'** is proved as follows. We know that the elements of $M_{XX}(\zeta)$, $M_{XZ}(\zeta)$, $M_{ZX}(\zeta)$, and $M_{ZZ}(\zeta)$ are polynomial functions of $\underline{\omega}$. Now from Theorem B.1(2) we also know that the elements of $M_{XX}^{-1}(\zeta)$ are rational functions of the elements of $M_{XX}(\zeta)$. These facts imply that the elements of the matrix $T^{-1}L\{I_r \otimes A[M(\zeta)]\}L'$, where $A[M(\zeta)] = M_{ZZ}(\zeta) - M_{ZX}(\zeta)M_{XX}^{-1}(\zeta)M_{XZ}(\zeta)$ are rational functions of $\underline{\omega}$. Since $\Lambda_2'[M(\zeta)] = tr[T^{-1}L\{I_r \otimes A[M(\zeta)]\}L']$ and the trace of a matrix is the sum of its diagonal elements $\Lambda_2'[M(\zeta)]$ has partial derivatives with respect to all the elements of $\underline{\omega}$.

We shall now prove that Λ_2' is concave on **M**, i.e., for $0 < \alpha < 1$ and M^a, $M^b \in$ **M**

$$\Lambda_2'(M) > (1 - \alpha)\Lambda_2'(M^a) + \alpha\Lambda_2'(M^b)$$

where $M = (1 - \alpha)M^a + \alpha M^b$. We shall use the following inequality from Fedorov (1972, p. 20):

If A_j has dimension nxm and B_j is a square positive definite matrix of order mxm ($j = 1, 2$) then

$$(1 - \alpha)A_1B_1^{-1}A_1' + \alpha A_2B_2^{-1}A_2'$$

$$> [(1 - \alpha)A_1 + \alpha A_2][(1 - \alpha)B_1 + \alpha B_2]^{-1}[(1 - \alpha)A_1' + \alpha A_2'] \quad ,$$

$$0 < \alpha < 1 \quad . \qquad (G.1)$$

We have that

$$M_{ZZ} - M_{ZX}M_{XX}^{-1}M_{XZ}$$

$$= (1 - \alpha)M_{ZZ}^a + \alpha M_{ZZ}^b - [(1 - \alpha)M_{ZX}^a + \alpha M_{ZX}^b]$$

$$[(1 - \alpha)M_{XX}^a + \alpha M_{XX}^b]^{-1}[(1 - \alpha)M_{XZ}^a + M_{XZ}^b]$$

$$> (1 - \alpha)M_{ZZ}^a + \alpha M_{ZZ}^b - (1 - \alpha)M_{ZX}^a(M_{XX}^a)^{-1}M_{XZ}^a - \alpha M_{ZX}^b(M_{XX}^b)^{-1}M_{XZ}^b$$

(applying inequality (G.1), with $A_1 = M_{ZX}^a$, $A_2 = M_{ZX}^b$,

$B_1 = M_{XX}^a$, $B_2 = M_{XX}^b$)

$$= (1 - \alpha)[M_{ZZ}^a - M_{ZX}^a(M_{XX}^a)^{-1}M_{XZ}^a] + \alpha[M_{ZZ}^b - M_{ZX}^b(M_{XX}^b)^{-1}M_{XZ}^b] \quad .$$

Hence

$$\mathrm{tr}[T^{-1}L\{I_r \otimes (M_{ZZ} - M_{ZX}M_{XX}^{-1}M_{XZ})\}L']$$

$$\geq (1 - \alpha)\mathrm{tr}[T^{-1}L\{I_r \otimes (M^a_{ZZ} - M^a_{ZX}(M^a_{XX})^{-1}M^a_{XZ})\}L']$$

$$+ \alpha\ \mathrm{tr}[T^{-1}L\{I_r \otimes (M^b_{ZZ} - M^b_{ZX}(M^b_{XX})^{-1}M^b_{XZ})\}L'] \quad .$$

This inequality implies that

$$\Lambda_2'[M] \geq (1 - \alpha)\Lambda_2'[M^a] + \alpha\Lambda_2'[M^b] \quad .$$

REFERENCES

Aitken, A. C. (1935). On least-squares and linear combination of observations. Proceedings of the Royal Society of Edinburgh 55, 42-48.

Aitken, A. C. (1951). Determinants and Matrices. New York: Oliver and Boyd Interscience Publishers, Inc.

Aitkinson, A. C. (1972). Planning experiments to detect inadequate regression models. Biometrika 59, 275-293.

Aitkinson, A. C. and Fedorov, V. V. (1975). Optimal design: experiments for discriminating between several models. Biometrika 62, 289-303.

Billingsley, P. (1978). Probability and Measure. New York: John Wiley and Sons.

Box, G. E. P., and Draper, N. R. (1959). A basis for the selection of response surface designs. JASA, 54, 622-654.

Box, G. E. P., and Draper, N. R. (1965). The Bayesian estimation of common parameters from several responses. Biometrika 52, 355-365.

Box, G. E. P., Hunter, W. G., MacGregor, J. F., and Erjavec, J. (1973). Some problems associated with the analysis of multiresponse data. Technometrics 15, 33-51.

Box, G. E. P., and Watson, G. S. (1962). Robustness to nonnormality of regression tests. Biometrika 49, 93-106.

Conlon, M. (1979). SEARCH: A Controlled Random Search Procedure for Function Minimization. Center for Instructional and Research Computing Activities, 411 Weil Hall, University of Florida.

Draper, N. R., and Smith, H. (1981). Applied Regression Analysis, second edition. New York: John Wiley.

Dykstra, O. (1971). The augmentation of experimental data to maximize $|X'X|$. Technometrics 13, 682-288.

Fedorov, V. V. (1972). Theory of Optimal Experiments. New York: Academic Press.

Gillespie, R. P. (1951). Partial Differentiation. New York: Interscience Publishers, Inc.

Graybill, F. A. (1969). Introduction to Matrices with Applications in Statistics. Wadsworth Publishing Company, Inc.

Jones, E. R., and Mitchell, T. J. (1978). Design criteria for detecting model inadequacy. Biometrika 65, 541-551.

Khuri, A. I. (1983). A test for lack of fit of a linear multiresponse model. Technical Report No. 198, Department of Statistics, University of Florida, Gainesville, Florida.

Khuri, A. I., and Myers, R. H. (1981). Design related robustness of tests in regression models. Commun. Statist.-Theor. Meth. A10(3), 223-235.

Kizer, O., Charvarie, C. Laguerie, C., and Cassimatis, D. (1978). Quadriatic model of the behavior of a fluidized bed reactor: catalytic oxidation of benzene to maleic anhydride. Canadian Journal of Chemical Engineering 56, 716-724.

Kiefer, J. (1959). Optimum experimental designs. J. R. Statist. Soc. B 21, 272-319.

Lancaster, P. (1969). Theory of Matrices, New York: Academic Press.

Mardia, K. V. (1971). The effect of nonnormality on some multivariate tests and robustness to nonnormality in the linear model. Biometrika 58, 105-121.

McLean, D. D., Pritchard, D. J. Bacon, D. W., and Downie, J. (1979). Singularities in multiresponse modeling. Technometrics 21, 291-298.

Myers, R. H. (1971). Response Surface Methodology. Allyn and Bacon, Inc., Boston.

Price, W. L. (1977). A controlled random search procedure for global optimisation. The Computer Journal 20, 367-370.

Roy, S. N., Gnanadesikan, R., and Srivastava, J. N. (1971). Analysis and Design of Certain Quantitative Multiresponse Experiments. New York: Pergamon Press.

Silvey, S. D. (1980). Optimal Design. University Press: Cambridge.

St. John, R. C., and Draper, N. R. (1975). D-Optimality for regression designs: a review. Technometrics 17, 15-23.

Wynn, H. P. (1970). The sequential generation of D-optimal experimental designs. Ann. Math. Statist. 41, 1655-1664.

Zellner, A. (1962). An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias. JASA 57, 348-368.

BIOGRAPHICAL SKETCH

Manel Cooray Wijesinha was born in Panadura, Sri Lanka, on April 13, 1951, to Justin and Harriet Cooray. She received a bachelor's degree in mathematics from the University of Sri Lanka, Colombo, in 1975. After two years of teaching undergraduate mathematics and statistics at the University of Sri Lanka, Colombo, she came to the United States to begin her graduate work at the University of Missouri, St. Louis, where she also worked as a teaching assistant in mathematics. In September 1978, she enrolled in the graduate mathematics program at the University of Florida and a year later she formally began her graduate work in statistics. She obtained a Master of Statistics degree in 1980. While at the University of Florida Manel has worked as a teaching assistant in mathematics, and also as a graduate assistant in statistics. Her duties included computer programming in connection with statistical analysis, statistical consulting, and teaching undergraduate mathematics.

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

A. K. Khuri

Andre I. Khuri, Chairman
Associate Professor of Statistics

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

John A. Cornell

John A. Cornell
Professor of Statistics

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Esam M. Ahmed
Professor of Food Science
and Human Nutrition

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Victor Chew
Associate Professor of Statistics

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Kenneth P. Portier
Assistant Professor of Statistics

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

A. K. Khuri

Andre I. Khuri, Chairman
Associate Professor of Statistics

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

John A. Cornell

John A. Cornell
Professor of Statistics

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Esam M. Ahmed
Professor of Food Science
and Human Nutrition

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Victor Chew
Associate Professor of Statistics

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Kenneth P. Portier
Assistant Professor of Statistics